Variable-Length Constrained Coding and Kraft Conditions: The Parity-Preserving Case

Ron M. Roth, *Fellow, IEEE* Paul H. Siegel, *Life Fellow, IEEE*

***Abstract*—Previous work by the authors on parity-preserving fixed-length constrained encoders is extended to the variable-length case. Parity-preserving variable-length encoders are formally defined, and, to this end, Kraft conditions are developed for the parity-preserving variable-length setting. Then, a necessary and sufficient condition is presented for the existence of deterministic parity-preserving variable-length encoders for a given constraint. Examples are provided that show that there are coding ratios where parity-preserving variable-length encoders exist, while fixed-length encoders do not.**

***Index Terms*—Constrained codes, Kraft inequality, Parity-preserving encoders, Variable-length encoders.**

I. INTRODUCTION

In mass storage platforms, such as magnetic and optical disks, user data is mapped (encoded) to binary sequences that satisfy certain combinatorial constraints. One common example of such a constraint is the (d, k)-runlength-limited (RLL) constraint, where the runs of 0's in a sequence are limited to have lengths at least d (to avoid inter-symbol interference) and at most k (to allow clock resynchronization) [8]. In virtually all applications, the encoder takes the form of a finite state machine, where user data is broken into binary blocks, and each block is mapped, in a state-dependent manner, into a binary *codeword*, so that the concatenation of the generated codewords satisfies the RLL constraint. In the case of *fixed-length encoders*, the input blocks all have the same length p, and the codewords all have the same length q, for prescribed positive integers p and q. The coding rate is then $p : q$.

In the mentioned storage applications, there is also a need to control the DC content of the recorded modulated sequence. One commonly used strategy to achieve DC control is allowing input blocks to be mapped to more than one codeword, and the encoder then selects the codeword that yields a better DC suppression [10, p. 29]. In the Blu-ray standard, this strategy is applied through the use of *parity-preserving encoders*: such encoders map each input block to a codeword that has the same parity (of the number of 1s), and DC control is achieved by reserving one bit in the input block to be set to a value that minimizes the DC contents [8, §11.4.3], [9], [11], [12], [13], [16].

This work was supported by Grants 2015816 and 2018048 from the United-States–Israel Binational Science Foundation (BSF), by NSF Grant CCF-BSF-1619053, and by Grants 1396/16 and 1713/20 from the Israel Science Foundation. This work, under the title "On parity-preserving variable-length constrained coding," was presented in part at the IEEE Int'l Symposium on Information Theory (ISIT), June 2020.

Ron M. Roth is with the Computer Science Department, Technion, Haifa 320003, Israel. This work was done in part while R.M. Roth was visiting the Center for Memory and Recording Research (CMRR), UC San Diego. Email: ronny@cs.technion.ac.il

Paul H. Siegel is with the ECE Department and CMRR, UC San Diego, La Jolla, CA 92023, USA. Email: psiegel@ucsd.edu

Most constructions of parity-preserving encoders that were proposed for commercial use were obtained by ad-hoc methods. In [14], we initiated a study of bi-modal encoders (which include parity-preserving encoders as a special case), focusing on fixed-length encoders; we will summarize the concepts that pertain to the fixed-length case, along with the main results of [14], as part of the background that we provide in Section II below. On the other hand, the existing ad-hoc parity-preserving constructions typically have *variable length*, where the length p of the input block and the length q of the respective codeword may depend on the encoder state, as well as on the input sequence (the *coding ratio*, p/q, nevertheless, is still fixed).

In this work, we present several results on parity-preserving variable-length encoders (in short, parity-preserving VLEs), focusing on deterministic encoders. To put our results into perspective, we mention that even in the ordinary setting (where parity preservation is not required), the known tools for analyzing and synthesizing VLEs are much less developed, compared to the fixed-length case. A summary of relevant (and mostly known) results on (ordinary) VLEs is provided in Section III. In Sections IV–V we turn to the parity-preserving setting. Much of the discussion in those sections deals in fact with the definition of parity-preserving VLEs, as it entails a (nontrivial) extension of the known Kraft conditions on variable-length coding to the parity-preserving case. This extension, which may be of independent interest, is developed in Section IV, followed in Section V by our main result, which is a necessary and sufficient condition for the existence of parity-preserving VLEs that are deterministic. We present several examples that demonstrate the advantages that parity-preserving VLEs may have over their fixed-length counterparts, in terms of the attainable coding ratios and encoding–decoding complexity.

II. FIXED-LENGTH GRAPHS AND ENCODERS

In this section, we extract from [10, Chapters 2–5] several basic definitions and properties pertaining to ordinary (namely, fixed-length) graphs and fixed-length encoders. We then quote the main result of [14], which applies, in particular, to parity-preserving fixed-length encoders.

A. Graphs and constraints

A (finite labeled directed ordinary) graph is a graph $G = (V, E, L)$ where V is a nonempty finite set of states, E is a finite set of edges, and $L : E \to \Sigma$ is an edge labeling. We say that a (finite) word $\boldsymbol{w}$ over Σ is *generated* by a path π in G if $\boldsymbol{w}$ is obtained by reading the labels along π; the length of $\boldsymbol{w}$ then equals the length of π (being the number of edges along π). A graph G is *deterministic* if no two outgoing edges from the same state in G have the same label. A deterministic graph is a special case of a graph with *finite anticipation*: the anticipation of a graph G is the smallest integer $a \ge 0$ (if any) such that any two paths with the same initial state that generate the same word of length $a{+}1$ must have the same initial edge (a deterministic graph corresponds to the case where the anticipation is 0). Having finite anticipation, in turn, implies (generally) that the graph is *lossless*: no two paths with the same initial state and the same terminal state generate the same word.

A *constraint* S over an alphabet Σ is the set of all words that are generated by paths in a graph G; we then say that G *presents* S and write $S = S(G)$. Every constraint S can be presented by a deterministic graph. The *capacity* of S is defined by $\mathsf{cap}(S) = \lim_{\ell\to\infty}(1/\ell)\log_2|S \cap \Sigma^\ell|$ (where, by sub-additivity, the limit indeed exists). It is known that $\mathsf{cap}(S) = \log_2 \lambda(A_G)$ where $\lambda(A_G)$ denotes the spectral radius (Perron eigenvalue) of the adjacency matrix A_G of any lossless (in particular, deterministic) presentation G of S.

A graph G is *irreducible* if it is strongly connected, namely, for any two states u and v in G there is a path from u to v. A constraint S is irreducible if it can be presented by a deterministic irreducible graph. For irreducible constraints, there is a unique deterministic graph presentation that has the smallest number of states; such a presentation is called the *Shannon cover* of S.

Example 1. Let S be the constraint over the alphabet $\Sigma = \{a, b, c, d\}$ which is presented by the graph G in Figure 1. The graph G is deterministic and irreducible (in fact, it is the Shannon cover of S). The adjacency matrix of A_G is given by

$$A_G = \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} ,$$

and $\lambda(A_G) = 2$, with a respective eigenvector $\boldsymbol{x} = (2\ \ 1)^\top$. Hence, $\mathsf{cap}(S) = \log_2 \lambda(A_G) = \log_2 2 = 1$. □

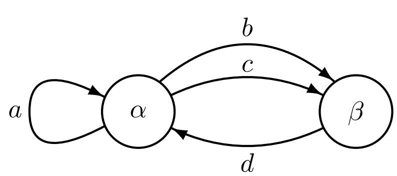

Fig. 1. Graph G for Example 1.

The power G^t of a graph $G = (V, E, L)$ is the graph with the same set of states V and edges that are the paths of length t in G; the label of an edge in G^t is the length-t word generated by the path. For $S = S(G)$ the power S^t is defined as $S(G^t)$.

B. Fixed-length encoders

Given a constraint S and a positive integer n, a *(fixed-length) (S, n)-encoder* is a lossless graph $\mathcal{E}$ such that $S(\mathcal{E}) \subseteq S$ and each state has out-degree n. An (S, n)-encoder exists if and only if $\log_2 n \le \mathsf{cap}(S)$. In a *tagged* (S, n)-encoder, each edge is assigned an input tag from a finite alphabet Υ of size n, such that edges outgoing from the same state have distinct tags. A tagged encoder is *(m, a)-sliding-block decodable* if all paths that generate a given word of length $m{+}a{+}1$ share the same tag on their $(m{+}1)$st edges.

A *(tagged) rate $p : q$ encoder* for a constraint S is a tagged $(S^q, 2^p)$-encoder (the tag alphabet Υ is then assumed to be $\{0, 1\}^p$); such an encoder exists if and only if $p/q \le \mathsf{cap}(S)$.

Given a square nonnegative integer matrix A and a positive integer n, an *(A, n)-approximate eigenvector* is a nonnegative nonzero integer vector $\boldsymbol{x}$ that satisfies the inequality $A\boldsymbol{x} \ge n\boldsymbol{x}$ componentwise. The set of all (A, n)-approximate eigenvectors will be denoted by $\mathcal{X}(A, n)$. Given a constraint S presented by a deterministic graph G and a positive integer n, the state-splitting algorithm provides a method for transforming G, through an (A_G, n)-approximate eigenvector, into an (S, n)-encoder with finite anticipation.

Example 2. Letting G and S be as in Example 1, the graph in Figure 2 is a tagged $(S, 2)$-encoder (or a rate $1 : 1$ encoder for S), where each edge is assigned a tag from $\{0, 1\}$ (the notation "s/w" next to an edge specifies the tag s and the label w of the edge). The encoder is obtained by splitting state α in G into two states: state α' inherits the outgoing edges labeled by b and c, and state α'' inherits the self-loop labeled a (this splitting is implied by the $(A_G, 2)$-approximate eigenvector $\boldsymbol{x} = (2\ \ 1)^\top$, which is also a true eigenvector of A_G, where state α in G is assigned a weight of 2, and state β has weight 1). The encoder is not deterministic, but it is $(0, 1)$-sliding-block decodable (and hence has anticipation 1): a label of an edge uniquely determines the initial state of the edge and, so, any word $\boldsymbol{w} \in S$ of length 2 uniquely determines the first edge of any path that generates $\boldsymbol{w}$. □

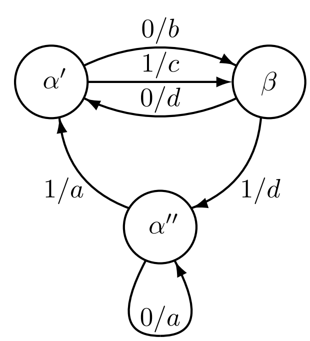

Fig. 2. Tagged fixed-length $(S, 2)$-encoder for Example 2.

C. Parity-preserving fixed-length encoders

Let Σ be an alphabet and fix a partition $\{\Sigma_0, \Sigma_1\}$ of Σ. The symbols in Σ_0 (respectively, Σ_1) will be referred to as the *even* (respectively, *odd*) symbols of Σ. Extending the definition of parity to words, we say that a word $\boldsymbol{w}$ over Σ is even (respectively, odd) if $\boldsymbol{w}$ contains an even (respectively, odd) number of symbols from Σ_1. The set of even (respectively, odd) words in Σ^t will be denoted by $(\Sigma^t)_0$ (respectively,

$(\Sigma^t)_1$). In the practical scenario where $\Sigma = \{0,1\}^p$, with Σ_0 and Σ_1 consisting of the binary p-tuples with even and odd parity, respectively (according to the common meaning of parity), a parity of a word in Σ^t, too, coincides with the ordinary meaning of this term.

Given a graph H with labeling in Σ, for $\mathsf{b} \in \{0,1\}$, we denote by H_{b} the subgraph of H containing only the edges with labels in Σ_{b}.

Example 3. Let $\Sigma = \{a,b,c,d\}$ and assume the partition $\{\Sigma_0, \Sigma_1\}$, where

$$\Sigma_0 = \{a,b\} \quad \text{and} \quad \Sigma_1 = \{c,d\}\ . \tag{1}$$

For the graph G in Figure 1, the subgraphs G_0 and G_1 with respect to this partition are shown in Figures 3 and 4. □

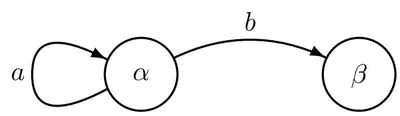

Fig. 3. Subgraph G_0 for Example 3.

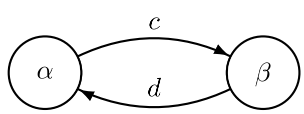

Fig. 4. Subgraph G_1 for Example 3.

Let S be a constraint over an alphabet Σ, fix a partition $\{\Sigma_0, \Sigma_1\}$ of Σ, and let n_0 and n_1 be positive integers. A *(fixed-length)* (S, n_0, n_1)*-encoder* $\mathcal{E}$ is an $(S, n_0{+}n_1)$-encoder such that for each $\mathsf{b} \in \{0,1\}$, the subgraph $\mathcal{E}_{\mathsf{b}}$ is an (S, n_{b})-encoder. A *rate* $p:q$ *parity-preserving (fixed-length) encoder* for S is a tagged $(S^q, 2^{p-1}, 2^{p-1})$-encoder in which the tag (in $\{0,1\}^p$) that is assigned to each edge has the same parity as the edge label (when seen as a word in Σ^q). Conversely, in any $(S^q, 2^{p-1}, 2^{p-1})$-encoder we can assign tags from $\{0,1\}^p$ to the edges so that the parities of the tags and the labels match on each edge.

Example 4. Letting Σ and S be as in Example 1, the $(S,2)$-encoder in Figure 2 is *not* an $(S,1,1)$-encoder with respect to the partition (1) of Σ, since both outgoing edges from state α' (respectively, state α'') have the same parity. In fact, using Theorem 1(a) below, it was shown in [14] that for the constraint S and for the partition (1), there is no $(S^t, 2^{t-1}, 2^{t-1})$-encoder for any positive integer t, namely, a coding ratio of 1 cannot be achieved by any parity-preserving (fixed-length) encoder, for any t. □

The next theorem follows from the results of [14] (see Theorem 1, Corollary 5, and §III-A therein).

Theorem 1 ([14])**.** *Let S be an irreducible constraint, presented by an irreducible deterministic graph G, and let n_0 and n_1 be positive integers. Then the following holds.*

(a) There exists an (S, n_0, n_1)-encoder, if and only if $\mathcal{X}(A_{G_0}, n_0) \cap \mathcal{X}(A_{G_1}, n_1) \neq \emptyset$.

(b) There exists a deterministic (S, n_0, n_1)-encoder, if and only if $\mathcal{X}(A_{G_0}, n_0) \cap \mathcal{X}(A_{G_1}, n_1)$ contains a 0–1 vector.

III. Variable-length graphs and encoders

In this section, we summarize several definitions and properties relating to variable-length graphs and variable-length encoders (see also [10, §6.4]).

A. Variable-length graphs

In a variable-length graph (in short, VLG), the labels of the edges may be words of any positive (finite) length over the label alphabet Σ; the length of the edge is then defined as the length of its label. Given a VLG H, the constraint $S(H)$ that is presented by H is defined as the set of all (consecutive) *sub-words* of words obtained by concatenating the labels that are read along finite paths in H. Equivalently, $S(H)$ is the constraint presented by the (ordinary) graph G obtained from H by replacing each length-ℓ edge e in H by a path of ℓ length-1 edges (connected through newly introduced dummy states) which generates the label of e. The *follower set* of a state u in H, denoted $\mathcal{F}_H(u)$, is the set of all *prefixes* of words that are generated by finite paths that start at u.

A VLG H is called deterministic if the labels of the outgoing edges from each state in H form a *prefix-free* list, namely, no label is a prefix of any other label. The notions of losslessness and irreducibility carry over from ordinary graphs: H is lossless if no two paths in H that share the same initial state and terminal state generate the same word, and it is irreducible if it is strongly connected.

Example 5. Letting G and S be as in Example 1, the VLG H in Figure 5 is irreducible and deterministic, and it presents S, i.e., $S(H) = S(G) = S$. In particular, we have $\mathcal{F}_H(\alpha) = \mathcal{F}_G(\alpha)$. □

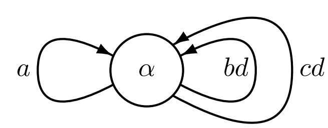

Fig. 5. VLG H for Example 5.

Remark 1. The follower-set equality, $\mathcal{F}_H(\alpha) = \mathcal{F}_G(\alpha)$, in Example 5 is in fact an instance of a more general property. Let S be an irreducible constraint and let G be its Shannon cover (i.e., the unique deterministic presentation of S with the smallest number of states). Also, let H be an irreducible deterministic VLG that presents S. We can obtain from H an ordinary irreducible deterministic graph H' (with length-1 edges) by transforming the outgoing edges from each state in H into a tree. From the uniqueness of the Shannon cover (and, specifically, from [10, Theorem 2.12(b)]) we get that the follower sets of the states of H' coincide with the follower sets of the states of G. Hence, for every state u in H there exists a state v in G such that $\mathcal{F}_H(u) = \mathcal{F}_G(v)$. □

B. Variable-length encoders

Let Υ be a finite alphabet[1] and let $\mathcal{L}$ be a finite list of nonempty finite words over Υ (the empty word is the unique word of length 0). We say that $\mathcal{L}$ is *exhaustive* if every word over Υ either has a prefix in $\mathcal{L}$ or is a prefix of some word in $\mathcal{L}$. The next result is well known [4, p. 298].

Theorem 2. *Given an alphabet Υ and a nonnegative integer sequence $\boldsymbol{\mu} = (\mu_\ell)_{\ell \geq 1}$ with finite support, there exists an exhaustive prefix-free list $\mathcal{L}$ over Υ such that*

$$\mu_\ell = \left| \mathcal{L} \cap \Upsilon^\ell \right| \ , \quad \ell = 1, 2, 3, \cdots \ ,$$

if and only if $\boldsymbol{\mu}$ satisfies the Kraft inequality with equality, namely:

$$\sum_{\ell \geq 1} \frac{\mu_\ell}{|\Upsilon|^\ell} = 1 \ . \tag{2}$$

Let S be a constraint over an alphabet Σ and let n be a positive integer. Also, let $\mathcal{E} = (V, E, L)$ be a VLG, and for every $u \in V$ and $\ell \geq 1$, denote by $\mu_\ell(u)$ the number of edges of length ℓ outgoing from u in $\mathcal{E}$. We say that $\mathcal{E}$ is a *variable-length* (S, n)*-encoder* (in short, an (S, n)*-VLE*) if the following conditions hold.

(E1) $\mathcal{E}$ is lossless,
(E2) $S(\mathcal{E}) \subseteq S$, and—
(E3) for every $u \in V$:

$$\sum_{\ell \geq 1} \frac{\mu_\ell(u)}{n^\ell} = 1 \ .$$

(This definition reduces to that of a fixed-length (S, n)-encoder when $\mu_\ell(u) = 0$ for every $u \in V$ and $\ell > 1$.)

Extending now the notion of tagging to the variable-length case, let Υ be a (base tag) alphabet of size $|\Upsilon| = n$. A tagging of an (S, n)-VLE $\mathcal{E}$ is an assignment of input tags—namely, words over Υ—to the edges of $\mathcal{E}$, such that:

(T1) the length of each input tag equals the length of (the label of) the edge, and—
(T2) the input tags of the outgoing edges from each state in $\mathcal{E}$ form an exhaustive prefix-free list over Υ.

Theorem 2 and condition (E3) guarantee that every (S, n)-VLE can be tagged consistently with conditions (T1)–(T2). Condition (T1) means that the *coding ratio* is fixed to be 1 at all edges, regardless of their length (as we argue in Remark 2 below, any fixed coding ratio can be reduced to the case of a coding ratio of 1). We note that this is the variable-length encoding model assumed in [1], [2], [6], and this model is more restrictive than the one in [7], where the coding ratio needs to be constant only along *cycles* in the encoder (see Figure 7 below).

Example 6. Letting Σ and S be as in Example 1, the graph H in Figure 5 is a deterministic $(S, 2)$-VLE. Taking $\Upsilon = \{0, 1\}$, one possible tag assignment to (the labels of) the edges of H is shown in Table I. The coding rate is $1 : 1$ when the input tag is 0, and $2 : 2$ when the input tag starts with a 1; namely, the coding ratio at each state is 1, so this encoder is capacity-achieving. Note that this tag assignment is parity-preserving with respect to the partition (1) of Σ. In contrast, recall from Example 4 that for this partition, a coding rate of $t : t$ cannot be achieved by any parity-preserving *fixed-length* encoder for S for any positive integer t. □

TABLE I
POSSIBLE TAG ASSIGNMENT FOR THE ENCODER IN FIGURE 5.

0	$\leftrightarrow$	a
10	$\leftrightarrow$	bd
11	$\leftrightarrow$	cd

Example 7. Letting Σ and S be as in Example 1, the graph $\mathcal{E}$ in Figure 6 presents another $(S, 2)$-VLE. The coding rate at state α' is $3 : 3$, as it has eight outgoing edges with labels in Σ^3, and the coding rate at α'' and at β is $2 : 2$, as each state has four outgoing edges labeled from Σ^2; the coding ratio at each state is therefore 1, making $\mathcal{E}$ capacity-achieving. However, $\mathcal{E}$ is not deterministic (there are two edges labeled bda and two labeled cda outgoing from state α', two edges labeled aa outgoing from α'', and two labeled da from state β). Nevertheless, $\mathcal{E}$ has finite anticipation and is therefore lossless: the first symbol of a label uniquely determines the length of the label as well as the initial state, and a label and the first symbol of the next label within a sequence uniquely determine the edge.

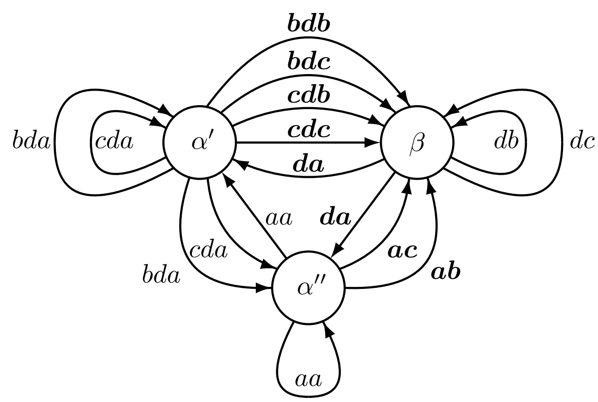

Fig. 6. VLE $\mathcal{E}$ for the constraint presented by Figure 1.

Consider now the following partition $\{\Sigma_0, \Sigma_1\}$ of Σ:

$$\Sigma_0 = \{a\} \quad \text{and} \quad \Sigma_1 = \{b, c, d\} \ . \tag{3}$$

With respect to this partition, the eight outgoing edges from state α' in $\mathcal{E}$ are equally divided between $(\Sigma^3)_0$ and $(\Sigma^3)_1$, and the four outgoing edges from each of the states α'' and β are equally divided between $(\Sigma^2)_0$ and $(\Sigma^2)_1$ (odd labels are marked in boldface in Figure 6). Hence, taking the tag alphabet $\Upsilon = \{0, 1\}$, we can achieve a coding ratio of 1 by a parity-preserving VLE. One possible parity-preserving tag assignment to the edges of $\mathcal{E}$ is shown in Table II. Similarly to the partition (1), it was shown in [14] that for the partition (3), too, one cannot achieve a coding ratio of 1 by any parity-preserving fixed-length encoder for S.

[1]We use here the notation Υ for an alphabet (instead of Σ) since in the context of variable-length encoders, that alphabet will be the alphabet of tags.

TABLE II
POSSIBLE TAG ASSIGNMENT FOR THE ENCODER IN FIGURE 6.

State α'		
$000, 011$	$\leftrightarrow$	bda
$101, 110$	$\leftrightarrow$	cda
001	$\leftrightarrow$	$\boldsymbol{bdb}$
010	$\leftrightarrow$	$\boldsymbol{bdc}$
100	$\leftrightarrow$	$\boldsymbol{cdb}$
111	$\leftrightarrow$	$\boldsymbol{cdc}$

State α''		
$00, 11$	$\leftrightarrow$	aa
01	$\leftrightarrow$	$\boldsymbol{ac}$
10	$\leftrightarrow$	$\boldsymbol{ab}$

State β		
$01, 10$	$\leftrightarrow$	$\boldsymbol{da}$
00	$\leftrightarrow$	db
11	$\leftrightarrow$	dc

The encoder in Figure 6 can be obtained from (an untagged copy of) the encoder in Figure 2 by replacing the outgoing edges from state α' with the eight paths of length 3 that start at that state and, similarly, replacing the outgoing edges from each of the states α'' and β with the four paths of length 2 that start at the state. $\square$

To summarize, for the constraint S of Example 1, Examples 6 and 7 present, respectively, (capacity-achieving) parity-preserving VLEs with a coding ratio of 1 for the two partitions (1) and (3): the first VLE is deterministic, while the other is not. In fact, we show in Appendix A that for the partition (3), one cannot achieve a coding ratio of 1 by any deterministic parity-preserving VLE (unless one uses a degenerate base tag alphabet containing only even symbols). On the other hand, there exists such an encoder under some relaxation of the notion of fixed coding ratio, following the encoding model considered in [7]: the tagged encoder $\mathcal{E}^\circ$ in Figure 7 maintains a coding ratio of 1 *along each cycle*. It

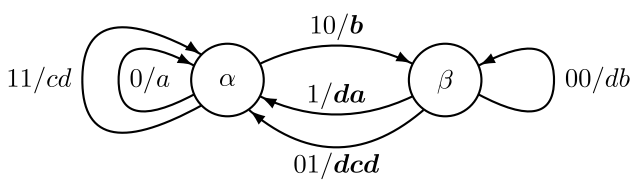

Fig. 7. Second VLE $\mathcal{E}^\circ$ for the constraint presented by Figure 1.

is easily seen that while at state α, each outgoing edge is uniquely determined by its first symbol, and while at state β, an outgoing edge is uniquely determined by its first two symbols.

Remark 2. Extending the terminology from fixed-length encoders, in a tagged VLE at a (fixed) coding ratio p/q for a constraint S, input tags are words over the (base) tag alphabet, and the length of a tag of each edge equals p/q times the edge length. The set of tags of the outgoing edges from each state must form an exhaustive prefix-free list. Assuming that $\gcd(p, q) = 1$, the length ℓ of an edge must be divisible by q, so we can consider the constraint S^q instead and regard each length-ℓ label over Σ as a word of length ℓ/q over Σ^q. Accordingly, we can group the $p\ell/q$ symbols in each tag into ℓ/q blocks of length p. Doing so, the coding ratio becomes 1. $\square$

Example 8. Let S be the $(2, \infty)$-RLL constraint, whose Shannon cover is given by the graph G in Figure 8. The

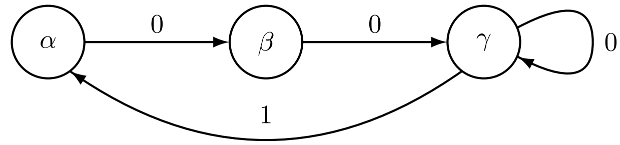

Fig. 8. Shannon cover G of the $(2, \infty)$-RLL constraint.

capacity of S is approximately 0.5515, so there exists a rate $1 : 2$ fixed-length encoder for S (namely, an $(S^2, 2)$-encoder); such a (tagged) encoder $\mathcal{E}$ is shown in Figure 9 (note that in this case, $S(\mathcal{E})$ is strictly contained in S^2). This encoder

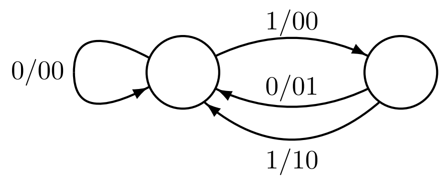

Fig. 9. Rate $1 : 2$ fixed-length encoder $\mathcal{E}$ for the $(2, \infty)$-RLL constraint.

is not deterministic; in fact, the smallest integer p for which there exists a rate $p : 2p$ deterministic fixed-length encoder for S is $p = 7$, as this is the smallest integer for which the set $\mathcal{X}(A_G^{2p}, 2^p)$ contains a 0–1 vector (see [10, Theorem 7.15]). Still, the encoder $\mathcal{E}$ is $(0, 1)$-sliding-block decodable.

On the other hand, the graph in Figure 10, with the tagging of Table III, is a deterministic VLE for S with a coding ratio of $1/2$ (see [6]; since the alphabet of S^2 consists of pairs of bits, we have used dots to delimit the symbols within each label). Note, however, that the tag assignment in Table III is *not* parity-preserving; we will return to this example in Examples 9 and 10 below. $\square$

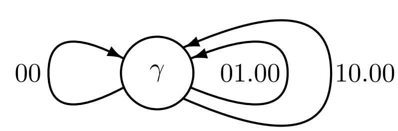

Fig. 10. VLE for the $(2, \infty)$-RLL constraint.

TABLE III
POSSIBLE TAG ASSIGNMENT FOR THE ENCODER IN FIGURE 10.

0	$\leftrightarrow$	00
10	$\leftrightarrow$	01.00
11	$\leftrightarrow$	10.00

C. Deterministic variable-length encoders

In this section, we focus on VLEs which are deterministic, and quote a necessary and sufficient condition for having such encoders.

Let $H=(V,E,L)$ be a VLG whose labels are over a finite alphabet Σ and let n be a positive integer. Fix some nonempty subset $V' \subseteq V$, and let $H'=(V',E',L')$ be the subgraph of H that is induced by V' (namely, E' consists of all the edges in H both of whose endpoints are in V'). For every $u \in V'$ and $\ell \ge 1$, denote by $\mu_\ell(u|V')$ the number of outgoing edges of length ℓ from u in H'. We say that V' is a set of *principal states in* H with respect to n if for every $u \in V'$:

$$\sum_{\ell \ge 1} \frac{\mu_\ell(u|V')}{n^\ell} \ge 1 \; . \tag{4}$$

It readily follows from this definition that V' is a set of principal states in a VLG H with respect to n, if and only if it is also so in the subgraph H' of H that is induced by V'.

The following result is essentially known (see [2], [5], [6]).

Theorem 3. *Let S be an irreducible constraint and let n and r be positive integers. There exists a deterministic (S,n)-VLE whose edges all have length at most r, if and only if S is presented by an irreducible deterministic VLG $H=(V,E,L)$ whose edges all have length at most r, and V contains a subset of principal states with respect to n.*[2]

We include a proof of the theorem both for completeness and for reference in our upcoming extension of this result to the parity-preserving case.

Proof of Theorem 3. Sufficiency follows by first looking at the subgraph H' of H that is induced by a set of principal states V'. We then (possibly) remove outgoing edges from states in H', starting with the longest outgoing edge and proceeding (if necessary) with edges in descending order of their lengths, until the inequality (4) becomes an equality at each state $u \in V'$.

To show necessity, suppose that $\mathcal{E}$ is a deterministic (S,n)-VLE. By shifting to an irreducible sink[3] of $\mathcal{E}$, we can assume that $\mathcal{E}$ is irreducible. Let G be the Shannon cover of S. By transforming the outgoing edges from each state in $\mathcal{E}$ into a tree (as in Remark 1), we get from [10, Lemma 2.13] that for every state Z in $\mathcal{E}$ there is a state u in G such that $\mathcal{F}_{\mathcal{E}}(Z) \subseteq \mathcal{F}_G(u)$. Let V' denote the states in G whose follower sets contain follower sets of states of $\mathcal{E}$; clearly, V' is not empty. For every $u \in V'$, let $Z(u)$ be some particular state Z in $\mathcal{E}$ such that $\mathcal{F}_{\mathcal{E}}(Z) \subseteq \mathcal{F}_G(u)$.

Next, we construct a deterministic VLG $H^*=(V',E^*,L^*)$ in which the outgoing edges from each state $u \in V'$ are defined as follows: for each edge $Z(u) \stackrel{\boldsymbol{w}}{\to} \tilde{Z}$ outgoing from $Z(u)$ in $\mathcal{E}$ (where $\tilde{Z}$ is the terminal state of the edge and $\boldsymbol{w}$ is its label), endow H^* with an edge $u \stackrel{\boldsymbol{w}}{\to} \tilde{u}$, where $\tilde{u}$ is the terminal state of the (unique) path in G that starts at u and generates the word $\boldsymbol{w}$; note that $\mathcal{F}_{\mathcal{E}}(\tilde{Z}) \subseteq \mathcal{F}_G(\tilde{u})$ and, therefore, $\tilde{u} \in V'$. By the construction it follows that H^* is deterministic and, by possibly shifting to an irreducible sink of H^*, we can assume that H^* is irreducible. One can easily show by induction on ℓ that all length-ℓ words in $\mathcal{F}_{H^*}(u)$ are contained in $\mathcal{F}_G(u)$, for every $u \in V'$; hence, $\mathcal{F}_{H^*}(u) \subseteq \mathcal{F}_G(u)$ for every $u \in V'$ and, in particular, $S(H^*) \subseteq S(G) = S$. Moreover, denoting by $\mu^*_\ell(u)$ the number of outgoing edges of length ℓ from state u in H^*, we have, for every $u \in V'$:

$$\sum_{\ell \ge 1} \frac{\mu^*_\ell(u)}{n^\ell} = 1 \; . \tag{5}$$

Thus, H^* is an irreducible deterministic (S,n)-VLE. Moreover, the length of each edge in H^* is at most the length of the longest edge in $\mathcal{E}$.

Next, we construct a VLG $H=(V,E,L)$ that contains H^* as a subgraph, as follows. Start with $(V,E,L) \leftarrow (V',E^*,L^*)$. Then, for each state $u \in V'$, let $\mathsf{r}(u)$ be the length of the longest edge outgoing from u in H^*. For every word $\boldsymbol{w} \in \mathcal{F}_G(u) \setminus \mathcal{F}_{H^*}(u)$ of length $\mathsf{r}(u)$ that does not have any prefix that labels any of the outgoing edges from u in H^*, endow H with an edge labeled $\boldsymbol{w}$ from u to the terminal state v of the path from u in G that generates $\boldsymbol{w}$ (in particular, insert v into V if it is not there already). Finally, iteratively endow H with the (length-1) outgoing edges (in G) from each state $u \in V \setminus V'$ (and insert their terminal states to V if they are not there already), until no new edges are added.

We claim that H is irreducible. Indeed, the subgraph H^* is irreducible, and every state $u \in V \setminus V'$ is reachable from V' in H (or else it would not have been inserted into V). Moreover, from each state $u \in V \setminus V'$ we can reach some state in V' in H by following the shortest path from u to V' in the Shannon cover G.

Secondly, we claim that H is deterministic. Indeed, at each state $u \in V'$ we only add edges of length $\mathsf{r}(u)$ whose labels do not have prefixes that label the existing outgoing edges from u in H^*, and at each state $u \in V \setminus V'$, the outgoing neighborhood from u in H is the same as that in G.

Thirdly, we claim that $\mathcal{F}_H(u) = \mathcal{F}_G(u)$ for every $u \in V$. We prove this by induction, showing that $\mathcal{F}_H(u) \cap \Sigma^\ell = \mathcal{F}_G(u) \cap \Sigma^\ell$ for every $\ell \ge 0$. The induction base $\ell = 0$ is trivial, due to the empty word. As for the induction step, the case $u \in V \setminus V'$ is immediate, while the case $u \in V'$ follows from the addition of the edges labeled by words $\boldsymbol{w} \in (\mathcal{F}_G(u) \setminus \mathcal{F}_{H^*}(u)) \cap \Sigma^{\mathsf{r}(u)}$ (whose prefixes do not label outgoing edges from u) to H. Irreducibility of both G and H and the equality $\mathcal{F}_H(u) = \mathcal{F}_G(u)$ (for some state $u \in V$) then imply that $S(H) = S(G)$.

Finally, since H^* is a subgraph of the subgraph H' of H that is induced by V', we get from (5) that (4) holds for every $u \in V'$, namely, V' is a set of principal states in H with respect to n. $\Box$

Remark 3. It follows from Remark 1 that when H is (irreducible, deterministic, and) reduced, its set of states is in effect a subset of the set of states of the Shannon cover G of $S(H)$. Therefore, any principal set of states V' of such an H consists of states of the Shannon cover of $S(H)$. $\Box$

Remark 4. It follows from the proof of the "if" part of Theorem 3 that if an irreducible deterministic VLG H contains

[2] Moreover, the graph H can be assumed to be *reduced*, namely, the follower sets of its states are distinct. For the case where all the edge lengths are 1, the graph H is the Shannon cover of S.

[3] An irreducible sink of $\mathcal{E}$ is an irreducible subgraph $\overline{\mathcal{E}} = (\overline{V}, \overline{E}, \overline{L})$ of $\mathcal{E}$ such that all the outgoing edges from $\overline{V}$ in $\mathcal{E}$ terminate in $\overline{V}$. Every graph has at least one irreducible sink [10, §2.5.1]. It is straightforward to see that an irreducible sink of an (S,n)-VLE is also an (S,n)-VLE.

a set V' of principal states with respect to n, then there is a deterministic (S,n)-VLE $\mathcal{E} = (V', \tilde{E}, \tilde{L})$ which is a subgraph of the subgraph H' of H that is induced by V'. Moreover, V' can be assumed to be the set of states of $\mathcal{E}$ (although $\mathcal{E}$ could then be reducible). $\square$

Given an ordinary irreducible deterministic graph G (with length-1 edges) and positive integers n and r, Franaszek described in [6] a polynomial-time algorithm for testing whether $S(G)$ can be presented by a VLG H that satisfies the conditions of Theorem 3 (see also [2], [3]). His algorithm, which is based on dynamic programming, effectively finds a set of principal states V' (which is a subset of the states of G) and a subgraph H' of H that is induced by V' (the graph H itself is not explicitly constructed in [6]).

Example 9. Let S be the $(2,\infty)$-RLL constraint, which is presented by the graph G in Figure 8, and take $n = 2$. Since there are no deterministic $(S^2, 2)$-encoders, we cannot have any principal states when $r = 1$.

Selecting $r = 2$, an application of Franaszek's algorithm from [6] to G^2 yields a (unique) set of principal states V' consisting only of state γ. Since without loss of generality H is reduced, that implies a unique subgraph H' that is induced by V', which is the graph in Figure 10 (see [6, §V]). $\square$

IV. PARITY-PRESERVING KRAFT CONDITIONS

In Section V, we provide a formal definition of a parity-preserving variable-length encoder. A key ingredient in that definition will be an adaptation of Theorem 2 to the parity-preserving case, which we do next; that adaptation may be of independent interest, beyond its use in this work. The main result of this section is Theorem 4 below, whose statement uses the following definition and notation.

Let Υ be a finite alphabet and assume a partition $\{\Upsilon_0, \Upsilon_1\}$ of Υ. Given a finite list $\mathcal{L}$ of nonempty words over Υ, the *(parity-preserving) length distribution* of $\mathcal{L}$ is a pair of nonnegative integer sequences $(\boldsymbol{\eta}{=}(\eta_\ell)_{\ell\ge1}, \boldsymbol{\omega}{=}(\omega_\ell)_{\ell\ge1})$, where

$$\eta_\ell = \left|\mathcal{L} \cap (\Upsilon^\ell)_0\right| \ \text{ and } \ \omega_\ell = \left|\mathcal{L} \cap (\Upsilon^\ell)_1\right| \ , \quad \ell = 1, 2, 3, \cdots .$$

In words, η_ℓ (respectively, ω_ℓ) is the number of even (respectively, odd) length-ℓ words in $\mathcal{L}$.

Given integers n and $\ell > 0$ and an integer sequence $\boldsymbol{\mu} = (\mu_i)_{i\ge1}$ with finite support, we define the following functional:

$$\mathsf{K}_\ell(\boldsymbol{\mu}, n) = n^\ell - \sum_{i=1}^{\ell} \mu_i \cdot n^{\ell-i} .$$

Given now positive integers n_0, n_1, and ℓ and a pair $(\boldsymbol{\eta}, \boldsymbol{\omega})$ of nonnegative integer sequences, each with finite support, define

$$\mathsf{K}_\ell^+ = \mathsf{K}_\ell(\boldsymbol{\eta} + \boldsymbol{\omega}, n_0 + n_1)$$

and

$$\mathsf{K}_\ell^- = \mathsf{K}_\ell(\boldsymbol{\eta} - \boldsymbol{\omega}, n_0 - n_1) .$$

Thus,

$$\begin{aligned} \mathsf{K}_\ell^\pm &= \mathsf{K}_\ell(\boldsymbol{\eta} \pm \boldsymbol{\omega}, n_0 \pm n_1) \\ &= (n_0 \pm n_1)^\ell - \sum_{i=1}^{\ell} (\eta_i \pm \omega_i)(n_0 \pm n_1)^{\ell-i} \\ &= (n_0 \pm n_1)^\ell \cdot \left(1 - \sum_{i=1}^{\ell} \frac{\eta_i \pm \omega_i}{(n_0 \pm n_1)^i}\right) , \end{aligned} \tag{6}$$

where the last equality applies for K_ℓ^- only when $n_0 \neq n_1$; when $n_0 = n_1$ we have instead:

$$\mathsf{K}_\ell^- = \omega_\ell - \eta_\ell . \tag{7}$$

Denoting hereafter by $\mathsf{r} = \mathsf{r}(\boldsymbol{\eta}, \boldsymbol{\omega})$ the largest index in the union of the supports of $\boldsymbol{\eta}$ and $\boldsymbol{\omega}$, the notation $\mathsf{K}^\pm = \mathsf{K}(\boldsymbol{\eta} \pm \boldsymbol{\omega}, n_0 \pm n_1)$ will stand for[4] $\mathsf{K}_\mathsf{r}^\pm$. Thus, (2) becomes

$$\mathsf{K}^+ = \mathsf{K}^+(\boldsymbol{\eta} + \boldsymbol{\mu}, n_0 + n_1) = 0 , \tag{8}$$

where we have taken $n_0 = |\Upsilon_0|$ and $n_1 = |\Upsilon_1|$.

The next theorem provides a necessary and sufficient condition for a pair $(\boldsymbol{\eta}, \boldsymbol{\omega})$ to be a (parity-preserving) length distribution of an exhaustive prefix-free list.

Theorem 4. *Given a partition $\{\Upsilon_0, \Upsilon_1\}$ of a finite alphabet Υ with $|\Upsilon_0| = n_0$ and $|\Upsilon_1| = n_1$, let $(\boldsymbol{\eta}, \boldsymbol{\omega})$ be a pair of nonnegative integer sequences, each with finite support. Then there exists an exhaustive prefix-free list over Υ with a length distribution $(\boldsymbol{\eta}, \boldsymbol{\omega})$, if and only if the following conditions hold.*

(a) $\mathsf{K}^+ = 0$, *and—*

(b) for every $\ell \ge 1$:

$$\mathsf{K}_\ell^+ \ge \left|\mathsf{K}_\ell^-\right| . \tag{9}$$

Remark 5. For $\ell \ge \mathsf{r} = \mathsf{r}(\boldsymbol{\eta}, \boldsymbol{\omega})$ we have $\mathsf{K}_\ell^\pm = (n_0 \pm n_1)^{\ell-\mathsf{r}} K^\pm$; hence, condition (a) is equivalent to requiring that $\mathsf{K}_\ell^+ = 0$ for *any* $\ell \ge \mathsf{r}$. Conditioning on (a), the inequality (9) for $\ell = \mathsf{r}$ is equivalent to

$$\mathsf{K}^+ = \mathsf{K}^- = 0 , \tag{10}$$

so it suffices to state condition (b) only for $1 \le \ell \le \mathsf{r}$: for larger ℓ, the inequality (9) follows from (10) (and holds with equality). $\square$

We prove Theorem 4 through a sequence of intermediary results, starting with the following equivalent formulation of conditions (a) and (b) (which is somewhat more explicit).

Lemma 5. *Conditions (a) and (b) in Theorem 4 are equivalent to the following conditions.*

(i) $\displaystyle\sum_{\ell\ge1} \frac{\eta_\ell + \omega_\ell}{(n_0+n_1)^\ell} = 1$,

(ii) $\displaystyle\sum_{\ell\ge1} \frac{\eta_\ell - \omega_\ell}{(n_0-n_1)^\ell} = 1$, *whenever $n_0 \neq n_1$, and—*

(iii) for every $\ell \ge 1$:

$$\sum_{i\ge1} \frac{\eta_{\ell+i} + \omega_{\ell+i}}{(n_0+n_1)^i} \ge \begin{cases} |\eta_\ell - \omega_\ell| & \text{if } n_0 = n_1 \\ \left|\displaystyle\sum_{i\ge1} \frac{\eta_{\ell+i} - \omega_{\ell+i}}{(n_0-n_1)^i}\right| & \text{if } n_0 \neq n_1 \end{cases} . \tag{11}$$

[4]There is a slight abuse in the notation $\mathsf{K}(\boldsymbol{\eta}-\boldsymbol{\omega}, n_0-n_1)$, since sometimes $\mathsf{r}(\boldsymbol{\eta}, \boldsymbol{\omega})$ is not uniquely determined from $\boldsymbol{\eta} - \boldsymbol{\omega}$.

Proof. Clearly, conditions (a) and (i) are equivalent. Next, we observe that for $n_0 = n_1$, the inequality (11) implies that $\eta_{\mathsf{r}} = \omega_{\mathsf{r}}$ for $\mathsf{r} = \mathsf{r}(\boldsymbol{\eta}, \boldsymbol{\omega})$. Hence, the following restatement of condition (ii) does not effectively change conditions (i)—(iii):

$$\textit{(ii')} \quad \begin{cases} \omega_{\mathsf{r}} - \eta_{\mathsf{r}} = 0 & \text{if } n_0 = n_1 \\ \displaystyle\sum_{\ell \geq 1} \frac{\eta_\ell - \omega_\ell}{(n_0 - n_1)^\ell} = 1 & \text{if } n_0 \neq n_1 \end{cases} .$$

By (6) and (7) it follows that conditions (i) and (ii') are equivalent to requiring $\mathsf{K}^+ = \mathsf{K}^- = 0$. Moreover, conditioning on (i) and (ii') (or conditioning on $0 = \mathsf{K}^+ \geq |\mathsf{K}^-|$), we have

$$\sum_{i=1}^{\ell} \frac{\eta_i \pm \omega_i}{(n_0 \pm n_1)^i} + \sum_{i \geq 1} \frac{\eta_{\ell+i} \pm \omega_{\ell+i}}{(n_0 \pm n_1)^{i+\ell}} = 1 .$$

Therefore,

$$\begin{aligned} \sum_{i \geq 1} \frac{\eta_{\ell+i} \pm \omega_{\ell+i}}{(n_0 \pm n_1)^i} &= (n_0 \pm n_1)^\ell \left(1 - \sum_{i=1}^{\ell} \frac{\eta_i \pm \omega_i}{(n_0 \pm n_1)^i} \right) \\ &= \mathsf{K}_\ell^\pm , \end{aligned}$$

and, so, (11) is equivalent to

$$\mathsf{K}_\ell^+ \geq \left| \mathsf{K}_\ell^- \right|$$

(even when $n_0 = n_1$). We conclude that conditions (i)–(iii) are equivalent to conditions (a)–(b). □

Lemma 6. *Given a partition $\{\Upsilon_0, \Upsilon_1\}$ of a finite alphabet Υ with $|\Upsilon_0| = n_0$ and $|\Upsilon_1| = n_1$, let $(\boldsymbol{\eta}, \boldsymbol{\omega})$ be a pair of nonnegative integer sequences, each with finite support. Then there exists an exhaustive prefix-free list over Υ with a length distribution $(\boldsymbol{\eta}, \boldsymbol{\omega})$, if and only if there exists a pair of nonnegative integer sequences $(\boldsymbol{y}{=}(y_\ell)_{\ell \geq 0}, \boldsymbol{z}{=}(z_\ell)_{\ell \geq 0})$ with finite support such that for every $\ell \geq 1$:*

$$\begin{aligned} \eta_\ell &= n_0 y_{\ell-1} + n_1 z_{\ell-1} - y_\ell \\ \omega_\ell &= n_1 y_{\ell-1} + n_0 z_{\ell-1} - z_\ell , \end{aligned} \tag{12}$$

where $y_0 \equiv 1$ and $z_0 \equiv 0$.

Proof. We start with proving the "only if" part. Let $(\boldsymbol{\eta}, \boldsymbol{\omega})$ be the length distribution of an exhaustive prefix-free list $\mathcal{L}$, and let $\mathcal{P}$ denote the set of words over Υ which are *proper* prefixes of words in $\mathcal{L}$; namely, a word $\boldsymbol{w}$ is in $\mathcal{P}$ if and only if there exists a nonempty word $\boldsymbol{w}'$ over Υ such that $\boldsymbol{w}\boldsymbol{w}' \in \mathcal{L}$ (in particular, $\mathcal{P}$ always contains the empty word). Since $\mathcal{L}$ is prefix-free, it cannot contain any of the (not necessarily proper) prefixes of the words in $\mathcal{P}$; in particular, $\mathcal{L} \cap \mathcal{P} = \emptyset$. On the other hand, since $\mathcal{L}$ is exhaustive, for any $s \in \Upsilon$ and $\boldsymbol{w} \in \mathcal{P}$, either $\boldsymbol{w}s \in \mathcal{L}$ or $\boldsymbol{w}s \in \mathcal{P}$ (but not both). Hence,

$$\{\boldsymbol{w}s \,:\, s \in \Upsilon, \boldsymbol{w} \in \mathcal{P}\} = \mathcal{L} \cup \mathcal{P}$$

and, so, for every $\ell \geq 1$ and $\mathsf{b} \in \{0, 1\}$:

$$\{\boldsymbol{w}s \,:\, s \in \Upsilon, \boldsymbol{w} \in \mathcal{P}\} \cap (\Upsilon^\ell)_{\mathsf{b}} = \big(\mathcal{L} \cap (\Upsilon^\ell)_{\mathsf{b}}\big) \cup \big(\mathcal{P} \cap (\Upsilon^\ell)_{\mathsf{b}}\big) . \tag{13}$$

For every $\ell \geq 0$, let y_ℓ (respectively, z_ℓ) denote the number of length-ℓ even (respectively, odd) words in $\mathcal{P}$:

$$\begin{aligned} y_\ell &= \left| \mathcal{P} \cap (\Upsilon^\ell)_0 \right| \\ z_\ell &= \left| \mathcal{P} \cap (\Upsilon^\ell)_1 \right| , \end{aligned}$$

where $y_0 = 1$ and $z_0 = 0$ (corresponding to the empty word, which is even). From (13) we then get:

$$\begin{aligned} n_0 y_{\ell-1} + n_1 z_{\ell-1} &= \eta_\ell + y_\ell \\ n_1 y_{\ell-1} + n_0 z_{\ell-1} &= \omega_\ell + z_\ell , \end{aligned}$$

thereby completing the proof of the "only if" part.

Next, we turn to proving the "if" part by induction on the value of $\mathsf{r} = \mathsf{r}(\boldsymbol{\eta}, \boldsymbol{\omega})$. We assume that (12) holds for some pair $(\boldsymbol{y}, \boldsymbol{z})$ with finite support, and we let r^* be the largest index in the union of the supports of $\boldsymbol{y}$ and $\boldsymbol{z}$. It follows from (12) that $\mathsf{r} = \mathsf{r}^* + 1$, i.e., $y_\ell = z_\ell = 0$ for $\ell \geq \mathsf{r}$. For the induction base $\mathsf{r} = 1$ we have $\mathsf{r}^* = 0$ and, so, $\eta_1 = n_0$ and $\omega_1 = n_1$, corresponding to $\mathcal{L} = \Upsilon$.

Suppose now that $\mathsf{r} > 1$ and define pairs $(\boldsymbol{\eta}', \boldsymbol{\omega}')$ and $(\boldsymbol{y}', \boldsymbol{z}')$ as follows:

$$\eta'_\ell = \begin{cases} \eta_\ell & \text{if } \ell < \mathsf{r} - 1 \\ \eta_{\mathsf{r}-1} + y_{\mathsf{r}-1} & \text{if } \ell = \mathsf{r} - 1 \\ 0 & \text{if } \ell > \mathsf{r} - 1 \end{cases} ,$$

$$\omega'_\ell = \begin{cases} \omega_\ell & \text{if } \ell < \mathsf{r} - 1 \\ \omega_{\mathsf{r}-1} + z_{\mathsf{r}-1} & \text{if } \ell = \mathsf{r} - 1 \\ 0 & \text{if } \ell > \mathsf{r} - 1 \end{cases} ,$$

and

$$y'_\ell = \begin{cases} y_\ell & \text{if } \ell \neq \mathsf{r} - 1 \\ 0 & \text{if } \ell = \mathsf{r} - 1 \end{cases} , \quad z'_\ell = \begin{cases} z_\ell & \text{if } \ell \neq \mathsf{r} - 1 \\ 0 & \text{if } \ell = \mathsf{r} - 1 \end{cases} .$$

It can be easily verified that those pairs satisfy (12), namely, for every $\ell \geq 1$:

$$\begin{aligned} \eta'_\ell &= n_0 y'_{\ell-1} + n_1 z'_{\ell-1} - y'_\ell \\ \omega'_\ell &= n_1 y'_{\ell-1} + n_0 z'_{\ell-1} - z'_\ell . \end{aligned}$$

Moreover, $\mathsf{r}(\boldsymbol{\eta}', \boldsymbol{\omega}') < \mathsf{r} = \mathsf{r}(\boldsymbol{\eta}, \boldsymbol{\omega})$. Hence, by the induction hypothesis, there exists an exhaustive prefix-free list $\mathcal{L}'$ whose length distribution is $(\boldsymbol{\eta}', \boldsymbol{\omega}')$. We construct from $\mathcal{L}'$ a new list $\mathcal{L}$ as follows. We select a subset $\mathcal{P}_{\mathsf{r}-1} \subseteq \mathcal{L}' \cap \Upsilon^{\mathsf{r}-1}$ consisting of $y_{\mathsf{r}-1}$ arbitrary words out of the $\eta'_{\mathsf{r}-1} = \eta_{\mathsf{r}-1} + y_{\mathsf{r}-1}$ words in $\mathcal{L}' \cap (\Upsilon^{\mathsf{r}-1})_0$, and $z_{\mathsf{r}-1}$ additional words out of the $\omega'_{\mathsf{r}-1} = \omega_{\mathsf{r}-1} + z_{\mathsf{r}-1}$ words in $\mathcal{L}' \cap (\Upsilon^{\mathsf{r}-1})_1$. We then replace each word $\boldsymbol{w} \in \mathcal{P}_{\mathsf{r}-1}$ by the $n_0 + n_1$ words $\boldsymbol{w}s$, where $s \in \Upsilon$, i.e.,

$$\mathcal{L} = (\mathcal{L}' \setminus \mathcal{P}_{\mathsf{r}-1}) \cup \{\boldsymbol{w}s \,:\, s \in \Upsilon, \boldsymbol{w} \in \mathcal{P}_{\mathsf{r}-1}\} .$$

The list $\mathcal{L}$ is both exhaustive and prefix-free, and it satisfies:

$$\left| \mathcal{L} \cap (\Upsilon^\ell)_0 \right| = \begin{cases} \eta'_\ell & \text{if } \ell < \mathsf{r} - 1 \\ \eta'_{\mathsf{r}-1} - y_{\mathsf{r}-1} & \text{if } \ell = \mathsf{r} - 1 \\ n_0 y_{\mathsf{r}-1} + n_1 z_{\mathsf{r}-1} & \text{if } \ell = \mathsf{r} \\ 0 & \text{if } \ell > \mathsf{r} \end{cases} ,$$

namely, $\left| \mathcal{L} \cap (\Upsilon^\ell)_0 \right| = \eta_\ell$ for all $\ell \geq 1$. In a similar way we also have $\left| \mathcal{L} \cap (\Upsilon^\ell)_1 \right| = \omega_\ell$, thereby completing the proof of the "if" part. □

Remark 6. From (12) we get

$$\begin{aligned} \sum_{\ell \geq 1} \frac{\eta_\ell + \omega_\ell}{(n_0 + n_1)^\ell} &= \sum_{\ell \geq 1} \frac{(n_0{+}n_1)(y_{\ell-1}{+}z_{\ell-1}) - (y_\ell{+}z_\ell)}{(n_0 + n_1)^\ell} \\ &= \sum_{\ell \geq 1} \left(\frac{y_{\ell-1} + z_{\ell-1}}{(n_0 + n_1)^{\ell-1}} - \frac{y_\ell + z_\ell}{(n_0 + n_1)^\ell} \right) \\ &= 1 , \end{aligned}$$

consistently with (2) (or with (8)). $\square$

Lemma 7. *Given positive integers n_0 and n_1, let $(\boldsymbol{\eta}, \boldsymbol{\omega})$ be a pair of nonnegative integer sequences, each with finite support. Then (12) is satisfied by a unique pair of real sequences $(\boldsymbol{y}{=}(y_\ell)_{\ell \ge 0}, \boldsymbol{z}{=}(z_\ell)_{\ell \ge 0})$ of finite support, and the values y_ℓ and z_ℓ are determined for every $\ell \ge 1$ by (the unique solution for (y_ℓ, z_ℓ) of) the following two equations:*

$$y_\ell + z_\ell = \sum_{i \ge 1} \frac{\eta_{\ell+i} + \omega_{\ell+i}}{(n_0 + n_1)^i} \tag{14}$$

and

$$y_\ell - z_\ell = \begin{cases} \omega_\ell - \eta_\ell & \text{if } n_0 = n_1 \\ \displaystyle\sum_{i \ge 1} \frac{\eta_{\ell+i} - \omega_{\ell+i}}{(n_0 - n_1)^i} & \text{if } n_0 \ne n_1 \end{cases} . \tag{15}$$

Proof. Replacing ℓ by $\ell + 1$ in (12) and then adding (respectively, subtracting) the two equations in (12), we obtain:

$$\eta_{\ell+1} \pm \omega_{\ell+1} = (n_0 \pm n_1)(y_\ell \pm z_\ell) - (y_{\ell+1} \pm z_{\ell+1}) .$$

This, in turn, yields the following backward recurrence for the values of $y_\ell \pm z_\ell$ (where we assume that $n_0 \ne n_1$ in the recurrence for $y_\ell - z_\ell$):

$$y_\ell \pm z_\ell = \frac{\eta_{\ell+1} \pm \omega_{\ell+1}}{n_0 \pm n_1} + \frac{y_{\ell+1} \pm z_{\ell+1}}{n_0 \pm n_1} .$$

Finally, we get (14) and (15) by repeated substitution, assuming the initial condition $y_\ell = z_\ell = 0$ for any sufficiently large $\ell \ge \mathsf{r}(\boldsymbol{\eta}, \boldsymbol{\omega})$. When $n_0 = n_1$, we get (15) directly simply by subtracting the two equations in (12). $\square$

Corollary 8. *Using the notation of Lemma 7, the pair $(\boldsymbol{y}, \boldsymbol{z})$ satisfies (12) for $\ell = 1$ with $(y_0, z_0) = (1, 0)$, if and only if*

$$\sum_{\ell \ge 1} \frac{\eta_\ell + \omega_\ell}{(n_0 + n_1)^\ell} = 1 \tag{16}$$

and (when $n_0 \ne n_1$)

$$\sum_{\ell \ge 1} \frac{\eta_\ell - \omega_\ell}{(n_0 - n_1)^\ell} = 1 . \tag{17}$$

Proof. The conditions (16)–(17) are equivalent to requiring that (14)–(15) be consistent with the initial condition $(y_0, z_0) = (1, 0)$ for $\ell = 0$. $\square$

Remark 7. The conditions on $(\boldsymbol{\eta}, \boldsymbol{\omega})$ in Lemma 7, combined with (16)–(17), guarantee that the solutions (y_ℓ, z_ℓ) of (14)–(15) are integer pairs for every $\ell \ge 1$; this can be seen if—instead of using (14)–(15)—we compute (y_ℓ, z_ℓ) iteratively for $\ell = 1, 2, 3, \cdots$, using the following recurrences (which are implied by (12)),

$$\begin{aligned} y_\ell &= n_0 y_{\ell-1} + n_1 z_{\ell-1} - \eta_\ell \\ z_\ell &= n_1 y_{\ell-1} + n_0 z_{\ell-1} - \omega_\ell , \end{aligned}$$

along with the initial condition $(y_0, z_0) = (1, 0)$. $\square$

Proof of Theorem 4. By Lemma 7, Corollary 8, and Remark 7, conditions (i) and (ii) in Lemma 5 are necessary and sufficient for having a pair of integer sequences $(\boldsymbol{y}, \boldsymbol{z})$ with $(y_0, z_0) = (1, 0)$ that satisfies (12). By Lemma 6, it remains to show that condition (iii) in Lemma 5 is necessary and sufficient for these sequences to be also nonnegative. Indeed, y_ℓ and z_ℓ are nonnegative if and only if

$$y_\ell + z_\ell \ge |y_\ell - z_\ell| ,$$

which, by (14)–(15), is equivalent to (11). $\square$

As we pointed out in Remark 5, the equality $\mathsf{K}^+ = \mathsf{K}^- = 0$ is equivalent to condition (a) in Theorem 4 combined with the requirement that the inequality (9) holds for all $\ell \ge \mathsf{r} = \mathsf{r}(\boldsymbol{\eta}, \boldsymbol{\mu})$. One may wonder if the remaining $\mathsf{r} - 1$ inequalities in condition (b) are independent in the sense that, conditioning on $\mathsf{K}^+ = \mathsf{K}^- = 0$, no subset of them implies the rest. In Appendix B, we show that this in fact holds, with the exception of the case $n_0 = n_1 = 1$.

V. Parity-preserving variable-length encoders

In this section, we provide a formal definition of a parity-preserving variable-length encoder. We then state a necessary and sufficient condition for having a parity-preserving VLE which is deterministic.

A. Definition of parity-preserving variable-length encoders

Let S be a constraint over an alphabet Σ and assume a partition $\{\Sigma_0, \Sigma_1\}$ of Σ. Also, let $\mathcal{E} = (V, E, L)$ be a VLG, and for every $u \in V$ and $\ell \ge 1$, denote by $\eta_\ell(u)$ (respectively, $\omega_\ell(u)$) the number of edges of length ℓ outgoing from u in $\mathcal{E}$ that have even (respectively, odd) labels (when the labels are regarded as words over Σ). Writing

$$\boldsymbol{\eta}(u) = (\eta_\ell(u))_{\ell \ge 1} \quad \text{and} \quad \boldsymbol{\omega}(u) = (\omega_\ell(u))_{\ell \ge 1} ,$$

the pair $(\boldsymbol{\eta}(u), \boldsymbol{\omega}(u))$ thus stands for the length distribution of the set of labels of the outgoing edges from u in H.

Fix now n_0 and n_1 to be positive integers, and for every $u \in V$ define

$$\mathsf{K}_\ell^\pm(u) = \mathsf{K}_\ell(\boldsymbol{\eta}(u) \pm \boldsymbol{\omega}(u), n_0 \pm n_1)$$

and

$$\mathsf{K}^\pm(u) = \mathsf{K}_\mathsf{r}(\boldsymbol{\eta}(u) \pm \boldsymbol{\omega}(u), n_0 \pm n_1) ,$$

where $\mathsf{r} = \mathsf{r}(u) = \mathsf{r}(\boldsymbol{\eta}(u), \boldsymbol{\omega}(u))$. We say that $\mathcal{E}$ is a *(parity-preserving) (S, n_0, n_1)-VLE* if for every $u \in V$ it satisfies the three conditions (E1)–(E3) in Section III-B, as well as the following fourth condition:

(E4) for every $\ell \ge 1$:

$$\mathsf{K}_\ell^+(u) \ge \left|\mathsf{K}_\ell^-(u)\right| .$$

(We note that condition (E3) can be rewritten as:

(E3) $\mathsf{K}^+(u) = 0$

and, so, by (E4) we also have $\mathsf{K}^-(u) = 0$.)

Now, let Υ be a base tag alphabet of size $n_0 + n_1$ that has a partition $\{\Upsilon_0, \Upsilon_1\}$ with $|\Upsilon_0| = n_0$ and $|\Upsilon_1| = n_1$. A (parity-preserving) tagging of an (S, n_0, n_1)-VLE is an assignment of input tags to the edges of $\mathcal{E}$ such that conditions (T1)–(T2) in Section III-B hold, and, in addition:

(T3) at each edge, the parity of the input tag (as a word over Υ) is the same as the parity of the label (as a word over Σ).

It follows from Theorem 4 and conditions (E3)–(E4) that every (S, n_0, n_1)-VLE can be tagged consistently with (T3).

B. Deterministic parity-preserving variable-length encoders

The main result of this section is Theorem 9 below, which is the parity-preserving counterpart of Theorem 3: it presents a necessary and sufficient condition for having a deterministic parity-preserving VLE.

Let Σ be an alphabet which is partitioned into $\{\Sigma_0, \Sigma_1\}$ and let $H = (V, E, L)$ be a VLG whose labels are over Σ. Fix some nonempty subset $V' \subseteq V$ and positive integers n_0 and n_1, and for every $u \in V'$ and $\ell \geq 1$, let $(\boldsymbol{\eta}(u|V'), \boldsymbol{\omega}(u|V'))$ be the length distribution of the set of labels of the outgoing edges from u in the subgraph $H' = (V', E', L')$ of H that is induced by V'. Also, for the purposes of this section, redefine

$$\mathsf{K}^{\pm}_{\ell}(u) = \mathsf{K}_{\ell}\left(\boldsymbol{\eta}(u|V') \pm \boldsymbol{\omega}(u|V'), n_0 \pm n_1\right)$$

and

$$\mathsf{K}^{\pm}(u) = \mathsf{K}_{\mathsf{r}}\left(\boldsymbol{\eta}(u|V') \pm \boldsymbol{\omega}(u|V'), n_0 \pm n_1\right) \ ,$$

where $\mathsf{r} = \mathsf{r}(u) = \mathsf{r}(\boldsymbol{\eta}(u|V'), \boldsymbol{\omega}(u|V'))$.[5] We say that V' is a set of *(parity-preserving) principal states in* H with respect to (n_0, n_1) if for every $u \in V'$:

$$\mathsf{K}^{+}(u) \leq -\left|\mathsf{K}^{-}(u)\right| \tag{18}$$

and

$$\mathsf{K}^{+}_{\ell}(u) \geq \left|\mathsf{K}^{-}_{\ell}(u)\right| \ , \quad \ell = 1, 2, \ldots, \mathsf{r}(u) - 1 \ . \tag{19}$$

Clearly, V' is a set of principal states in a VLG H (with respect to (n_0, n_1)), if and only if it is also so in the subgraph H' of H that is induced by V'.

For the special case where H is a deterministic (S, n_0, n_1)-VLE, conditions (E3)–(E4) imply that all the states of H form a set of principal states with respect to (n_0, n_1), with (18) replaced by the stronger condition

$$\mathsf{K}^{+}(u) = \mathsf{K}^{-}(u) = 0 \ . \tag{20}$$

Theorem 9. *Let S be an irreducible constraint over an alphabet Σ, assume a partition $\{\Sigma_0, \Sigma_1\}$ of Σ, and let n_0, n_1, and r be positive integers. There exists a deterministic (S, n_0, n_1)-VLE whose edges all have length at most r, if and only if S is presented by an irreducible deterministic VLG $H = (V, E, L)$ whose edges all have length at most r, and V contains a subset of principal states with respect to (n_0, n_1).*

Proof. The proof of the "only if" part builds upon the respective part in the proof of Theorem 3. Specifically, given a deterministic (S, n_0, n_1)-VLE $\mathcal{E}$, we define the set V' as in that proof and construct the VLE $H^* = (V', E^*, L^*)$. For every $u \in V'$, the length distribution of the set of labels of the outgoing edges from u in H^* is the same as the respective set for $Z(u)$ in $\mathcal{E}$. Hence, by conditions (E3)–(E4) it follows that H^* satisfies conditions (19) and (20). Then, when we form H from H^*, the change made at states $u \in V'$ is limited to adding outgoing edges of length $\mathsf{r}(u)$. Clearly, such a change has no effect on the terms appearing in (19). As for the terms in (19), let y^+ (respectively, y^-) be the number of even-labeled (respectively, odd-labeled) outgoing edges that were added to state u (all of which of length $\mathsf{r}(u)$). By (20) (when stated for H^*) we get that, in H,

$$\mathsf{K}^{+}(u) = -y^+ - y^- \quad \text{and} \quad \mathsf{K}^{-}(u) = -y^+ + y^- \ ,$$

thereby implying (18) (when stated for H, yet still with respect to the subset V').

Turning to the "if" part of the proof of Theorem 3, we need to show that we can remove edges from the subgraph H' of H that is induced by the set of principal states V' so that the resulting subgraph $\mathcal{E}$ satisfies (19) and (20). Fix some state $u \in V'$ in H', and suppose that we remove y^+ (respectively, y^-) even-labeled (respectively, odd-labeled) outgoing edges from state u, all of length $\mathsf{r} = \mathsf{r}(u)$. Similarly to what we had in the "only if" proof, such removal does not affect the terms in (19), yet it changes the values of $\mathsf{K}^+ = \mathsf{K}^+(u)$ and $\mathsf{K}^- = \mathsf{K}^-(u)$ into $\mathsf{K}^+ + y^+ + y^-$ and $\mathsf{K}^- + y^+ - y^-$, respectively; so, in order to satisfy (20), we require that y^+ and y^- be such that

$$\mathsf{K}^{\pm} + y^+ \pm y^- = 0 \ ,$$

namely,

$$y^{\pm} = -\frac{1}{2}(\mathsf{K}^{+} \pm \mathsf{K}^{-}) \ . \tag{21}$$

Noting that K^+ and K^- have the same parity, it follows that $y^{\pm}$ satisfying (21) are integers. Moreover, by condition (18) they are also nonnegative.

To complete the proof, it remains to show that there indeed exist $y^{\pm}$ edges that can be removed from H' at state u, namely, that $y^+ \leq \eta_{\mathsf{r}}$ and $y^- \leq \omega_{\mathsf{r}}$. Observing that

$$\mathsf{K}^{\pm} = (n_0 \pm n_1)\mathsf{K}^{\pm}_{\mathsf{r}-1} - (\eta_{\mathsf{r}} \pm \omega_{\mathsf{r}}) \ ,$$

we have:

$$\begin{aligned} y^{\pm} &\overset{(21)}{=} -\frac{1}{2}(\mathsf{K}^{+} \pm \mathsf{K}^{-}) \\ &= -\frac{1}{2}\Big((n_0 + n_1)\mathsf{K}^{+}_{\mathsf{r}-1} - (\eta_{\mathsf{r}} + \omega_{\mathsf{r}}) \\ &\qquad \pm (n_0 - n_1)\mathsf{K}^{-}_{\mathsf{r}-1} \mp (\eta_{\mathsf{r}} - \omega_{\mathsf{r}})\Big) \ . \end{aligned}$$

Hence,

$$\begin{aligned} y^+ &= \eta_{\mathsf{r}} - \frac{1}{2}\left((n_0 + n_1)\mathsf{K}^{+}_{\mathsf{r}-1} + (n_0 - n_1)\mathsf{K}^{-}_{\mathsf{r}-1}\right) \\ &\overset{(19)}{\leq} \eta_{\mathsf{r}} - \frac{1}{2}\left((n_0 + n_1)\left|\mathsf{K}^{-}_{\mathsf{r}-1}\right| + (n_0 - n_1)\mathsf{K}^{-}_{\mathsf{r}-1}\right) \\ &\leq \eta_{\mathsf{r}} \ , \end{aligned}$$

with the first (respectively, second) inequality holding with equality if and only if $\mathsf{K}^{+}_{\mathsf{r}-1} = \left|\mathsf{K}^{-}_{\mathsf{r}-1}\right|$ (respectively, $\mathsf{K}^{-}_{\mathsf{r}-1} = 0$); namely, $y^+ = \eta_{\mathsf{r}}$ if and only if $\mathsf{K}^{+}_{\mathsf{r}-1} = \mathsf{K}^{-}_{\mathsf{r}-1} = 0$. Similarly,

$$\begin{aligned} y^- &= \omega_{\mathsf{r}} - \frac{1}{2}\left((n_0 + n_1)\mathsf{K}^{+}_{\mathsf{r}-1} - (n_0 - n_1)\mathsf{K}^{-}_{\mathsf{r}-1}\right) \\ &\overset{(19)}{\leq} \omega_{\mathsf{r}} - \frac{1}{2}\left((n_0 + n_1)\left|\mathsf{K}^{-}_{\mathsf{r}-1}\right| - (n_0 - n_1)\mathsf{K}^{-}_{\mathsf{r}-1}\right) \\ &\leq \omega_{\mathsf{r}} \ , \end{aligned}$$

again, with $y^- = \omega_{\mathsf{r}}$ if and only if $\mathsf{K}^{+}_{\mathsf{r}-1} = \mathsf{K}^{-}_{\mathsf{r}-1} = 0$.

[5] That is, $\mathsf{K}^{\pm}_{\ell}(u)$, $\mathsf{K}^{\pm}(u)$, and $\mathsf{r}(u)$ are redefined here for the subgraph H' of H that is induced by the subset $V' \subseteq V$. For simplicity of notation, we have elected to make the dependence on V' only implicit, as V' will be understood from the context.

We conclude that conditions (18)–(19) guarantee that we can always remove edges from state $u \in V'$ in H' so that the resulting graph satisfies ((19) and) (20); note that this applies also to the case $y^+ = \eta_\mathsf{r}$ and $y^- = \omega_\mathsf{r}$, where the edge removal reduces the value of $\mathsf{r}(u)$, yet (20) will still hold since $\mathsf{K}^+_{\mathsf{r}-1} = \mathsf{K}^-_{\mathsf{r}-1} = 0$ (see Remark 5). $\square$

Example 10. Let S be the $(2, \infty)$-RLL constraint, which is presented by the graph G in Figure 8. Recall from Example 8 that there is no deterministic $(S^2, 2)$-encoder in this case and, so, there is no VLG H that satisfies the conditions of Theorem 3 for $r = 1$.

Turning to $r = 2$, recall from Example 9 that the VLE in Figure 10 is the unique induced subgraph H' of any (reduced) VLG H that satisfies the conditions of Theorem 3. Yet, assuming the ordinary definition of parity of binary words, the set of states $V' = \{\gamma\}$ of H' is not a set of (parity-preserving) principal states (in H' and therefore in H) with respect to $(n_0, n_1) = (1, 1)$. Hence, for $r = 2$, there is no deterministic $(S^2, 1, 1)$-VLE.

On the other hand, there exists a deterministic $(S^2, 1, 1)$-VLE for $r = 3$, as shown in Figure 11, along with the tag assignment in Table IV. This encoder is a subgraph of the

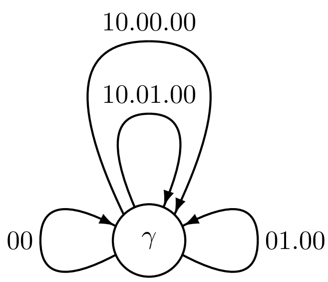

Fig. 11. Parity-preserving VLE for the $(2, \infty)$-RLL constraint.

TABLE IV
TAG ASSIGNMENT FOR THE ENCODER IN FIGURE 11.

0	$\leftrightarrow$	00
10	$\leftrightarrow$	01.00
110	$\leftrightarrow$	10.01.00
111	$\leftrightarrow$	10.00.00

VLG H shown in Figure 12, which is a deterministic VLG presentation of S^2 whose edges all have length at most 3, and $V' = \{\gamma\}$ is a set of principal states in H with respect to $(1, 1)$ (as such, H satisfies the conditions of Theorem 9).

Comparing to the fixed-length case, using Theorem 1(a), one can verify that there exists a (not necessarily deterministic) $(S^{2p}, 2^{p-1}, 2^{p-1})$-encoder, if and only if $p \geq 3$. For $p = 3$, any vector $\boldsymbol{x} \in \mathcal{X}(A_{(G^6)_0}, 4) \cap \mathcal{X}(A_{(G^6)_1}, 4)$ satisfies $\|\boldsymbol{x}\|_\infty \geq 6$ (and equality is attained only by $\boldsymbol{x} = (2\ 3\ 6)^\top$). By Corollaries 4 and 5 in [14] we then get that any rate $3 : 6$ parity-preserving fixed-length encoder for S must have at least six states and anticipation at least 2 (measured in 6-bit symbols); in contrast, recall that when there is no requirement for parity preservation, we have the simple encoder in Figure 9. Using Theorem 1(b),

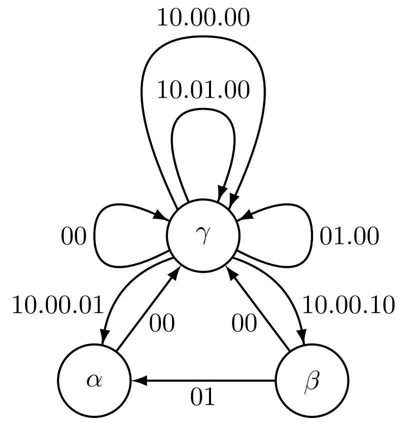

Fig. 12. Deterministic VLG presentation H of the second power of the $(2, \infty)$-RLL constraint.

one can determine that there exists a rate $p : 2p$ parity-preserving fixed-length encoder for S which is deterministic, (if and) only if $p \geq 8$. $\square$

Remark 8. Unlike Theorem 3, we do not have (as of yet) an extension of Franaszek's algorithm from [6] to the parity-preserving case; namely, a polynomial-time algorithm is yet to be found for determining whether, for given S, $\{\Sigma_0, \Sigma_1\}$, n_0, n_1, and r, there is a VLG H that satisfies the conditions of Theorem 9. (The problem, however, is still decidable, since there are only finitely many reduced VLGs H with edge lengths at most r such that $S(H) = S$.) $\square$

C. Discussion

In Appendix A, we show that for the constraint S of Example 1 and for the partition (3), there is no deterministic (S^t, n_0, n_1)-VLE, for any positive integers t, n_0, and n_1 such that $\log_2(n_0 + n_1) = t = \mathsf{cap}(S^t)$. In contrast, given any constraint $S = S(G)$ and positive integers n_0 and n_1 that satisfy the strict inequality $\log_2(n_0 + n_1) < \mathsf{cap}(S)$, it follows from (the proof of) Theorem 2 in [14] that, under mild conditions on the presentation G of S, there exist deterministic (fixed-length) $(S^r, n^{(r)}, n^{(r)})$-encoders H_r, $r = 1, 2, \ldots$, where $(\log_2 n^{(r)})/r \to \mathsf{cap}(S)$ $(> \log_2(n_0 + n_1))$ when $r \to \infty$. Thus, for sufficiently large r, each encoder H_r, when regarded as a VLG with all the edges having length r, contains a deterministic (S, n_0, n_1)-VLE as a subgraph.

As a sanity check, we next show that the states of H_r form a set of principal states with respect to (n_0, n_1). From $(\log_2 n^{(r)})/r > \log_2(n_0 + n_1)$ (for sufficiently large r) it follows that

$$(n_0 + n_1)^r + |n_0 - n_1|^r \leq 2n^{(r)} . \tag{22}$$

Now, the VLG H_r (whose edges all have length r) satisfies (19) vacuously (with V' taken as the whole set of states of H_r), and it also satisfies (18) since

$$\begin{aligned} \mathsf{K}^+(u) &= (n_0 + n_1)^r - \eta_r(u) - \omega_r(u) \\ &= (n_0 + n_1)^r - 2n^{(r)} \end{aligned}$$

$$
\begin{aligned}
&\overset{(22)}{\leq} && -\left|n_0 - n_1\right|^r \\
&= && -\left|(n_0 - n_1)^r - \eta_r(u) + \omega_r(u)\right| \\
&= && -\left|\mathsf{K}^-(u)\right| \ .
\end{aligned}
$$

We conclude that the states of H_r form a principal set of states and, so, by Theorem 9 there exists a deterministic $(S(H_r), n_0, n_1)$-VLE (and, as such, it is also an (S, n_0, n_1)-VLE).

When $G = (V, E, L)$ is an ordinary graph (whose edges all have length 1), condition (18) becomes, for every $u \in V'$:

$$
\begin{aligned}
(n_0 + n_1) - \sum_{v \in V'} (A_{G_0} + A_{G_1})_{u,v} & \\
\leq \quad & \left|(n_0 - n_1) - \sum_{v \in V'} (A_{G_0} - A_{G_1})_{u,v}\right| \ .
\end{aligned}
$$

This inequality can be rewritten as

$$
\sum_{v \in V'} (A_{G_0})_{u,v} \geq n_0 \quad \text{and} \quad \sum_{v \in V'} (A_{G_1})_{u,v} \geq n_1 \ ,
$$

and also as

$$
A_{G_0} \boldsymbol{x} \geq n_0 \boldsymbol{x} \quad \text{and} \quad A_{G_1} \boldsymbol{x} \geq n_1 \boldsymbol{x} \ ,
$$

where $\boldsymbol{x}$ is the 0–1 characteristic vector of the subset V' within V. Condition (19) becomes vacuous for ordinary graphs. It thus follows that a nonempty subset $V' \subseteq V$ is a set of principal states in G with respect to (n_0, n_1), if and only if its characteristic vector belongs to $\mathcal{X}(A_{G_0}, n_0) \cap \mathcal{X}(A_{G_1}, n_1)$. For G which is also deterministic, this coincides with Theorem 1(b).

Remark 9. When applying Theorems 3 and 9 to a finite-memory[6] constraint S and $r = 1$, the respective (fixed-length) deterministic (S, n_0, n_1)-encoder can be guaranteed to be also sliding-block decodable. On the other hand, when $r > 1$, edges in the encoder may have different lengths and, so, the output sequence consists of words (labels) of varying lengths over the alphabet Σ of S. State-independent decoding, however, should not assume the position of any given output symbol (of Σ) within the label (word) that it belongs to. This, in turn, imposes conditions beyond the Kraft conditions (19)–(20) on the lengths of the outgoing edges from each state in the encoder. When encoders do not have to be parity-preserving, such (sufficient) conditions were provided in [2] and [3]. Respective conditions are yet to be found for the parity-preserving case. □

In this paper, we focused mainly on parity-preserving VLEs which are deterministic. The study of the non-deterministic case is an open topic for future work. In particular, we can pose the following question: under what conditions can capacity be achieved (with equality) by parity-preserving VLEs? Recall that for the constraint S of Example 1 and for the partition (3), capacity cannot be achieved when the encoder is deterministic (as we show in Appendix A), nor when it is of fixed length (as we showed in [14]).

[6]A constraint S has finite memory if it can be defined through a finite list of forbidden words, i.e., $\boldsymbol{w} \in S$ if and only if $\boldsymbol{w}$ does not contain any word in that list as a sub-word [10, §2.3].

Appendix A
Nonexistence result for Example 7

Let Σ, G, and S be as in Example 1, and assume the partition (3) of Σ. We show that for this partition, there is no deterministic parity-preserving VLE at a coding ratio of 1. Specifically, we show that for every positive integers t, n_0, and n_1 such that $n_0 + n_1 = 2^t$, there is no deterministic parity-preserving (S^t, n_0, n_1)-VLE.

Suppose to the contrary that such a VLE exists, and let $\mathcal{E}$ be such an encoder with the smallest number of states. The encoder $\mathcal{E}$ is irreducible (or else its irreducible sink[7] would be a smaller encoder) and reduced (or else we could merge states with identical follower sets [10, §2.6.2]). By changing the outgoing edges from each state in $\mathcal{E}$ into a tree, we can get an (ordinary) irreducible deterministic graph G' (with edge labels of length 1 over Σ^t). The constraint $S(\mathcal{E}) = S(G')$ has capacity t, which is also the capacity of the (irreducible) constraint S^t in which it is contained. Hence, by [10, Problem 3.28] we have $S(\mathcal{E}) = S(G') = S^t = S(G^t)$ and, so, by Remark 1, for every state u in $\mathcal{E}$ there exists a state $v \in \{\alpha, \beta\}$ in G such that $\mathcal{F}_{\mathcal{E}}(u) = \mathcal{F}_{G^t}(v)$. It follows that $\mathcal{E}$ has no more than two states.

Assume first that $\mathcal{E}$ has only one state, in which case the edges in $\mathcal{E}$ are (variable-length) self-loops, corresponding to cycles in G^t. Note, however, that all the cycles in G (and, therefore, in G^t) generate even words, which means that any tagging of the edges of $\mathcal{E}$ forms a set $\mathcal{L}$ consisting only of even words over the base tag alphabet Υ. Yet, since we assume that both n_0 and n_1 are positive, the alphabet Υ contains at least one even symbol (say, 0) and one odd symbol (say, 1). But then, any word of the form $100\ldots0$ that is longer than the longest tag in $\mathcal{L}$ is neither a prefix of any tag in $\mathcal{L}$ (obviously), nor has it a prefix in $\mathcal{L}$; namely, $\mathcal{L}$ cannot be exhaustive.[8]

It remains to consider the case where $\mathcal{E}$ has two (inequivalent) states, which we denote by α and β to match their respective equivalent states in G. In fact, we will rule out the existence of a deterministic two-state $(S^t, 2^t)$-VLE, regardless of whether it is parity-preserving. Any deterministic $(S^t, 2^t)$-VLE, in turn, can be viewed as a deterministic $(S, 2)$-VLE, by regarding each length-ℓ label over Σ^t as a label of length $t\ell$ over Σ.

We recall the following definition of a parametrized adjacency matrix. Given a VLG $H = (V, E, L)$, for any two states $u, v \in V$ we denote by $\mu_\ell(u, v)$ the number of edges of length ℓ from u to v. For a positive real indeterminate θ, we define the parametrized adjacency matrix of H as the $|V| \times |V|$ matrix $A_H(\theta)$ whose entries are given by:

$$
(A_H(\theta))_{u,v} = \sum_{\ell \geq 1} \mu_\ell(u, v) \cdot \theta^{-\ell} \ .
$$

[7]Refer to Footnote 3 for the definition of an irreducible sink.

[8]The case $n_0 = 0$ can also be ruled out: the first label along any path that generates the (sufficiently long) even word $abdbd \ldots bd$ must end either with an a or with a d, and, so, that label must have odd length. On the other hand, any odd-length tag over an an all-odd alphabet cannot be even.

We point out that we can rule out an encoder with all-even labels also by using Lemma 5: it is easy to see that when $\omega_\ell = 0$ for all ℓ, conditions (i) and (ii') can hold simultaneously only when $n_1 = 0$.

We let $\theta_{\max}(H)$ denote the largest θ for which $\lambda(A_H(\theta)) = 1$. It is known that when H is lossless (in particular, deterministic), the capacity of $S(H)$ equals $\log_2 \theta_{\max}(H)$ [15, Theorem 1] (when all the edges have length 1 we have $A_H(\theta) = (1/\theta) \cdot A_H$, in which case $\theta_{\max}(H) = \lambda(A_H)$). It is also known that when H is irreducible, the mapping $\theta \mapsto \lambda(A_H(\theta))$ is strictly decreasing (and continuous) over $(0, \infty)$ [10, Proposition 3.12].

Turning now to the encoder $\mathcal{E}$, which we view as a deterministic two-state $(S, 2)$-VLE, we have

$$\begin{aligned}\sum_{v \in \{\alpha,\beta\}} (A_{\mathcal{E}})_{\beta,v}(2) &= \sum_{v \in \{\alpha,\beta\}} \sum_{\ell \geq 1} \mu_\ell(\beta, v) \cdot 2^{-\ell} \\ &= \sum_{\ell \geq 1} 2^{-\ell} \sum_{v \in \{\alpha,\beta\}} \mu_\ell(\beta, v) \\ &= 1 ,\end{aligned} \tag{23}$$

where the last equality follows from condition (E3). Denote by $\mathcal{B}$ the set of labels of the outgoing edges from state β in $\mathcal{E}$. By (23) we must have

$$|\mathcal{B}| = \sum_{\ell \geq 1} \sum_{v \in \{\alpha,\beta\}} \mu_\ell(\beta, v) \geq 2 .$$

Noting that all these labels start with the symbol d, we define $\mathcal{B}'$ to be the set of all words obtained by removing the leading symbol d from the words in $\mathcal{B}$. Since $\mathcal{E}$ is deterministic, the set $\mathcal{B}$ is prefix-free and, therefore, so is $\mathcal{B}'$. In particular, $\mathcal{B}'$ does not contain the empty word (since $|\mathcal{B}'| = |\mathcal{B}| \geq 2$). Next, construct from $\mathcal{E}$ an (irreducible) VLG $\mathcal{E}'$ by replacing the outgoing edges from state α with copies of the outgoing edges from state β, keeping the terminal states yet removing from each label its leading symbol d. Thus, for every $u, v \in \{\alpha, \beta\}$:

$$(A_{\mathcal{E}'}(\theta))_{u,v} = \begin{cases} \theta \cdot (A_{\mathcal{E}}(\theta))_{\beta,v} & \text{if } u = \alpha \\ (A_{\mathcal{E}}(\theta))_{\beta,v} & \text{if } u = \beta \end{cases} . \tag{24}$$

Since $\mathcal{B}'$ is prefix-free the graph $\mathcal{E}'$ is deterministic. Moreover, it can be easily verified that $\mathcal{F}_{\mathcal{E}'}(\alpha) \subseteq \mathcal{F}_G(\alpha)$ and, so, $S(\mathcal{E}') \subseteq S$. Hence, $\mathsf{cap}(S(\mathcal{E}')) \leq \mathsf{cap}(S) = 1$, which implies that

$$\theta_{\max}(\mathcal{E}') \leq 2 . \tag{25}$$

On the other hand, from (23)–(24) we get the following row sums in $A_{\mathcal{E}'}(2)$:

$$\sum_{v \in \{\alpha,\beta\}} (A_{\mathcal{E}'}(2))_{\alpha,v} = 2 \quad \text{and} \quad \sum_{v \in \{\alpha,\beta\}} (A_{\mathcal{E}'}(2))_{\beta,v} = 1 .$$

By [10, Proposition 3.13] we then get that $\lambda(A_{\mathcal{E}'}(2)) > 1$, i.e., $\theta_{\max}(\mathcal{E}') > 2$. Yet this contradicts (25).

Appendix B
Independence of the conditions in Theorem 4

Given positive integers n_0, n_1, and $r > 1$, we show that, unless $n_0 = n_1 = 1$, the inequalities (9) that correspond to $\ell = 1, 2, \ldots, r-1$ are independent (in the sense defined below) conditioned on $(\boldsymbol{\eta}, \boldsymbol{\omega})$ being a nonnegative integer pair with $\mathsf{r} = \mathsf{r}(\boldsymbol{\eta}, \boldsymbol{\omega}) = r$ that satisfies $\mathsf{K}^+ = \mathsf{K}^- = 0$. In particular, each of these inequalities is necessary, as it is not implied by the rest.

We introduce the following definition. Given positive integers n_0, n_1, and $r > 1$, a subset $\mathcal{Z} \subseteq \{1, 2, \ldots, r-1\}$ is said to be *admissible for* (n_0, n_1, r) if there exists a nonnegative integer pair $(\boldsymbol{\eta}, \boldsymbol{\omega})$ with $\mathsf{r} = \mathsf{r}(\boldsymbol{\eta}, \boldsymbol{\omega}) = r$ that satisfies $\mathsf{K}^+ = \mathsf{K}^- = 0$ yet violates (9) when (and only when) $\ell \in \mathcal{Z}$. The inequalities (9) are then said to be independent if every subset $\mathcal{Z} \subseteq \{1, 2, \ldots, r-1\}$ is admissible for (n_0, n_1, r).

We have the following lemma.

Lemma 10. *A subset $\mathcal{Z} \subseteq \{1, 2, \ldots, r-1\}$ is admissible for (n_0, n_1, r), if and only if there exists an integer pair $(\boldsymbol{y}{=}(y_\ell)_{\ell \geq 0}, \boldsymbol{z}{=}(z_\ell)_{\ell \geq 0})$ that satisfies the following conditions:*

- (C1) $y_\ell \leq n_0 y_{\ell-1} + n_1 z_{\ell-1}$ *and* $z_\ell \leq n_1 y_{\ell-1} + n_0 z_{\ell-1}$ *for every* $\ell \geq 1$,
- (C2) $y_{r-1} + z_{r-1} > 0$,
- (C3) $y_\ell = z_\ell = 0$ *when* $\ell \geq r$,
- (C4) $y_0 = 1$ *and* $z_0 = 0$, *and—*
- (C5) $\min\{y_\ell, z_\ell\} < 0$ *when (and only when)* $\ell \in \mathcal{Z}$.

Proof. We use (12) and (14)–(15) to define a one-to-one correspondence between integer pairs $(\boldsymbol{\eta}, \boldsymbol{\omega})$ and $(\boldsymbol{y}, \boldsymbol{z})$, both with finite support. Condition (C1) is equivalent to requiring that $\boldsymbol{\eta}$ and $\boldsymbol{\omega}$ are nonnegative, and conditions (C2)–(C3) are equivalent to having $\mathsf{r}(\boldsymbol{\eta}, \boldsymbol{\omega}) = r$ (and, when $n_0 = n_1$, also $\eta_r = \omega_r$). Conditioning on (C1)–(C3), we get by Corollary 8 that condition (C4) is equivalent to conditions (i) and (ii') in (the proof of) Lemma 5 being satisfied by $(\boldsymbol{\eta}, \boldsymbol{\omega})$; these conditions, in turn, are equivalent to requiring $\mathsf{K}^+ = \mathsf{K}^- = 0$.

Finally, conditioning on (C1)–(C4) (and, in particular, on on (i) and (ii')), we get from (14)–(15) that (11) (and, therefore, (9)) can be rewritten as

$$y_\ell + z_\ell \geq |y_\ell - z_\ell| ,$$

which, in turn, holds if and only if y_ℓ and z_ℓ are nonnegative. Hence, condition (C5) is equivalent to (9) being violated by $(\boldsymbol{\eta}, \boldsymbol{\omega})$ when (and only when) $\ell \in \mathcal{Z}$. $\square$

We now use Lemma 10 to identify the admissible subsets for any given (n_0, n_1, r). In particular, we show that when $\max\{n_0, n_1\} > 1$, every subset $\mathcal{Z} \subseteq \{1, 2, \ldots, r-1\}$ is admissible. We distinguish between three cases.

Case 1: $n_0 \geq n_1$ and $n_0 > 1$. We show that any subset $\mathcal{Z} \subseteq \{1, 2, \ldots, r-1\}$ is admissible for (n_0, n_1, r), for any $r > 1$. Indeed, given any such subset $\mathcal{Z}$, define the pair $(\boldsymbol{y}, \boldsymbol{z})$ by:

$$y_\ell = \begin{cases} 1 & \text{if } \ell = 0 \\ n_0 & \text{if } 1 \leq \ell < r \\ 0 & \text{if } \ell \geq r \end{cases}$$

and

$$z_\ell = \begin{cases} -1 & \text{if } \ell \in \mathcal{Z} \\ 0 & \text{otherwise} \end{cases} .$$

It can be readily checked that the pair $(\boldsymbol{y}, \boldsymbol{z})$ satisfies conditions (C1)–(C5) (where, for (C1)–(C2), we use the assumption that $n_0 > 1$).

Case 2: $n_1 \geq n_0$ and $n_1 > 1$. Here, too, any subset $\mathcal{Z} \subseteq \{1, 2, \ldots, r-1\}$ is admissible, for any $r > 1$: the proof is similar to Case 2, except that the pair $(\boldsymbol{y}, \boldsymbol{z})$ is now defined by

$$y_\ell = \begin{cases} 1 & \text{if } \ell = 0 \\ n_1 & \text{if } \ell \text{ is even and } 1 \le \ell < r \\ -1 & \text{if } \ell \text{ is odd and } \ell \in \mathcal{Z} \\ 0 & \text{otherwise} \end{cases}$$

and

$$z_\ell = \begin{cases} n_1 & \text{if } \ell \text{ is odd and } 1 \le \ell < r \\ -1 & \text{if } \ell \text{ is even and } \ell \in \mathcal{Z} \\ 0 & \text{otherwise} \end{cases} .$$

Case 3: $n_0 = n_1 = 1$. In this case, there are subsets of $\{1, 2, \ldots, r-1\}$ which are not admissible. For example, it can be verified that the inequality (9) for $\ell = 1$ is implied by $\mathsf{K}^+ = \mathsf{K}^- = 0$.

We next characterize the admissible subsets for $(1, 1, r)$. Given a subset $\mathcal{Z} \subseteq \{1, 2, \ldots, r-1\}$, define the integer sequence $\boldsymbol{\xi} = \boldsymbol{\xi}(\mathcal{Z}, r) = (\xi_1\, \xi_2 \ldots \xi_r)$ inductively as follows:

$$\xi_\ell = \begin{cases} 1 & \text{if } \ell = 1 \\ \xi_{\ell-1} - 1 & \text{if } \ell - 1 \in \mathcal{Z} \\ 2\xi_{\ell-1} & \text{otherwise} \end{cases} . \tag{26}$$

We have the following lemma.

Lemma 11. *A subset $\mathcal{Z} \subseteq \{1, 2, \ldots, r-1\}$ is admissible for $(1, 1, r)$, if and only if the sequence $\boldsymbol{\xi}$ as defined in (26) is all-positive.*

Proof. Starting with the "only if" part, suppose that there exists an integer pair $(\boldsymbol{y}, \boldsymbol{z})$ that satisfies conditions (C1)–(C5). Condition (C1) can be rewritten as

$$\max\{y_\ell, z_\ell\} \le y_{\ell-1} + z_{\ell-1} \tag{27}$$

which, with conditions (C4)–(C5), implies (by induction on ℓ) that $\max\{y_\ell, z_\ell\} \le \xi_\ell$ for every $\ell \in \{1, 2, \ldots, r-1\}$. In particular, for $\ell = r-1$ we have:

$$\begin{aligned} \xi_{r-1} &\ge \max\{y_{r-1}, z_{r-1}\} \\ &= y_{r-1} + z_{r-1} - \min\{y_{r-1}, z_{r-1}\} \\ &\overset{\text{(C2)}}{\ge} 1 - \min\{y_{r-1}, z_{r-1}\} \\ &\overset{\text{(C5)}}{\ge} \begin{cases} 2 & \text{if } r-1 \in \mathcal{Z} \\ 1 & \text{otherwise} \end{cases} , \end{aligned}$$

which, by (26), implies that $\xi_r > 0$. Moreover, by (26), the inequality $\xi_{r-1} > 0$ is possible only if $\xi_\ell > 0$ for every $\ell < r$.

Turning to the "if" part, given a sequence $\boldsymbol{\xi}$ as in (26) that is all-positive, we define the pair $(\boldsymbol{y}, \boldsymbol{z})$ as follows:

$$y_\ell = \begin{cases} 1 & \text{if } \ell = 0 \\ \xi_\ell & \text{if } 1 \le \ell < r \\ 0 & \text{otherwise} \end{cases}$$

and

$$z_\ell = \begin{cases} -1 & \text{if } \ell \in \mathcal{Z} \\ \xi_\ell & \text{if } \ell \in \{1, 2, \ldots, r-1\} \setminus \mathcal{Z} \\ 0 & \text{otherwise} \end{cases} .$$

Obviously, the pair $(\boldsymbol{y}, \boldsymbol{z})$ satisfies conditions (C3)–(C5). As for condition (C2), we have

$$\begin{aligned} y_{r-1} + z_{r-1} &= \begin{cases} \xi_{r-1} - 1 & \text{if } r-1 \in \mathcal{Z} \\ 2\xi_{r-1} & \text{otherwise} \end{cases} \\ &= \xi_r > 0 . \end{aligned}$$

Turning finally to condition (C1), the inequality (27) holds (trivially) with equality when $\ell = 1$ or when $\ell > r$, and is implied by condition (C2) when $\ell = r$. For the remaining range $\ell \in \{2, 3, \ldots, r-1\}$ we also have equality in (27), since:

$$\begin{aligned} \max\{y_\ell, z_\ell\} &= y_\ell = \xi_\ell \overset{(26)}{=} \begin{cases} \xi_{\ell-1} - 1 & \text{if } \ell - 1 \in \mathcal{Z} \\ 2\xi_{\ell-1} & \text{otherwise} \end{cases} \\ &= y_{\ell-1} + z_{\ell-1} . \end{aligned}$$

□

References

[1] R.L. Adler, J. Friedman, B. Kitchens, B.H. Marcus, "State splitting for variable-length graphs," *IEEE Trans. Inf. Theory,* 32 (1986), 108–113.

[2] M.-P. Béal, "The method of poles: A coding method for constrained channels," *IEEE Trans. Inf. Theory,* 36 (1990), 763–772.

[3] M.-P. Béal, "Extensions of the method of poles for code construction," *IEEE Trans. Inf. Theory,* 49 (2003), 1516–1523.

[4] R.E. Blahut, *Digital Transmission of Information,* Addison-Wesley, Reading, Massachusetts, 1990.

[5] P.A. Franaszek, "Sequence-state coding for digital transmission," *Bell Sys. Tech. J.,* 47 (1968), 143–155.

[6] P.A. Franaszek, "On synchronous variable length coding for discrete noiseless channels," *Inform. Control,* 15 (1969), 155–164.

[7] C.D. Heegard, B.H. Marcus, P.H. Siegel, "Variable-length state splitting with applications to average runlength-constrained (ARC) codes," *IEEE Trans. Inf. Theory,* 37 (1991), 759–777.

[8] K.A.S. Immink, *Codes for Mass Data Storage Systems,* Second Edition, Shannon Foundation Publishers, Eindhoven, The Netherlands, 2004.

[9] J.A.H.M. Kahlman, K.A.S. Immink, "Device for encoding/decoding N-bit source words into corresponding M-bit channel words, and vice versa," US Patent 5,477,222, 1995.

[10] B.H. Marcus, R.M. Roth, P.H. Siegel, *An Introduction to Coding for Constrained Systems,* Lecture Notes, 2001, available online at: ronny.cswp.cs.technion.ac.il/wp-content/uploads/sites/54/2016/05/chapters1-9.pdf

[11] T. Miyauchi, Y. Shinohara, Y. Iida, T. Watanabe, Y. Urakawa, H. Yamagishi, M. Noda, "Application of turbo codes to high-density optical disc storage using 17PP Code," *Jpn. J. Appl. Phys.,* 44 No. 5B (2005), 3471–3473.

[12] T. Narahara, S. Kobayashi, M. Hattori, Y. Shimpuku, G.J. van den Enden, J.A.H.M. Kahlman, M. van Dijk, R. van Woudenberg, "Optical disc system for digital video recording," *Jpn. J. Appl. Phys.,* 39 No. 2B (2000), 912–919.

[13] M. Noda, H. Yamagishi, "An 8-state DC-controllable run-length-limited code for the optical-storage channel," *Jpn. J. Appl. Phys.,* 44 No. 5B (2005), 3462–3466.

[14] R.M. Roth, P.H. Siegel, "On bi-modal constrained coding," *IEEE Trans. Inf. Theory,* 67 (2021), 1609–1621.

[15] C.E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.,* 27 (1948), 379–423 and 623–656.

[16] W.Y.H. Wilson, K.A.S. Immink, X.B. Xi, C.T. Chong, "A Comparison of two coding schemes for generating DC-free runlength-limited sequences," *Jpn. J. Appl. Phys.,* 39 No. 2B (2000), 815–818.

Ron M. Roth (Fellow, IEEE) received the B.Sc. degree in computer engineering, the M.Sc. in electrical engineering, and the D.Sc. in computer science from Technion—Israel Institute of Technology, Haifa, Israel, in 1980, 1984, and 1988, respectively. Since 1988 he has been with the Computer Science Department at Technion, where he now holds the General Yaakov Dori Chair in Engineering. During the academic years 1989–91 he was a Visiting Scientist at IBM Research Division, Almaden Research Center, San Jose, California, and during 1996–97, 2004–05, and 2011–2012 he was on sabbatical leave at Hewlett–Packard Laboratories, Palo Alto, California. He is the author of the book *Introduction to Coding Theory*, published by Cambridge University Press in 2006. Dr. Roth was an associate editor for coding theory in IEEE TRANSACTIONS ON INFORMATION THEORY from 1998 till 2001, and he is now serving as an associate editor in *SIAM Journal on Discrete Mathematics*. His research interests include coding theory, information theory, and their application to storage, computation, and the theory of complexity.

Paul H. Siegel (Life Fellow, IEEE) received the S.B. and Ph.D. degrees in mathematics from the Massachusetts Institute of Technology (MIT), Cambridge, MA, USA, in 1975 and 1979, respectively. He held a Chaim Weizmann Postdoctoral Fellowship with the Courant Institute, New York University, New York, NY, USA. He was with the IBM Research Division, San Jose, CA, USA, from 1980 to 1995. He joined the University of California at San Diego, CA, USA, in July 1995, as a Faculty Member, where he is currently a Distinguished Professor of electrical and computer engineering with the Jacobs School of Engineering. He is also affiliated with the Center for Memory and Recording Research, where he holds the Endowed Chair and served as the Director from 2000 to 2011. His research interests include information theory and communications, particularly coding and modulation techniques, with applications to digital data storage and transmission. He was a Member of the Board of Governors of the IEEE Information Theory Society from 1991 to 1996 and 2009 to 2014. He is a member of the National Academy of Engineering. He was the 2015 Padovani Lecturer of the IEEE Information Theory Society. He was a co-recipient of the 2007 Best Paper Award in signal processing and coding for data storage from the Data Storage Technical Committee of the IEEE Communications Society. He was a co-recipient of the 1992 IEEE Information Theory Society Paper Award and the 1993 IEEE Communications Society Leonard G. Abraham Prize Paper Award. He served as a Co-Guest Editor of the May 1991 Special Issue on Coding for Storage Devices of the IEEE TRANSACTIONS ON INFORMATION THEORY. He has served as an Associate Editor for Coding Techniques for the IEEE TRANSACTIONS ON INFORMATION THEORY from 1992 to 1995, and the Editor-in-Chief from July 2001 to July 2004. He was also a Co-Guest Editor of the May/September 2001 two-part issue on The Turbo Principle: From Theory to Practice and the February 2016 issue on Recent Advances in Capacity Approaching Codes of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS.